



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 147209

TO: Patricia Duffy
Location: REM-3B05/3C18
Art Unit: 1645
Monday, March 14, 2005
Case Serial Number: 09/674277

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1A69
Phone: 571-272-2518 *BOB*

barbara.obryen@uspto.gov

Search Notes

RUSH

1 10-13 1819, 20

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:06:25 ; Search time 8345.22 Seconds
(without alignment)
8645.645 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagcccgagatgaag.....ctggaagagcctggtcgac 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1489	100.0	1489	6	BD205211 Nucleotid
2	1489	100.0	1489	6	AX011297 Sequence
3	1476.4	99.2	92077	1	AF074613 Escherich
4	1476.4	99.2	92077	6	AX191727 Sequence
5	1476.4	99.2	92721	1	AB011549 Escherich
6	1476.4	99.2	92721	6	AX191725 Sequence
7	1081.4	72.6	2407	1	ECHECCTP
8	539.8	36.3	110000	1	Continuation (10 o
9	536.6	36.0	12268	1	AE013690 Yersinia
10	536.6	36.0	13632	1	AF135170 Yersinia
11	536.6	36.0	220050	1	AJ414156 Yersinia
12	536.6	36.0	290803	1	AE017128 Yersinia
13	394.6	26.5	295250	1	BX321862 Nitrosomo
14	384.8	25.8	2587	1	AB017595 Legionell
15	384.8	25.8	5249	1	AF276752 Legionell
16	384.8	25.8	110000	1	Continuation (3 of
17	383.2	25.7	110000	1	Continuation (3 of
18	382	25.7	11421	1	AE015517 Shewanell
19	380	25.5	110000	1	Continuation (3 of

C	20	368	24.7	1829	1	ECIS91TP
	21	357.6	24.0	2238	6	AR098264 Sequence
	22	357.6	24.0	2238	6	AR215252 Sequence
	23	356	23.9	290628	1	AE016983 Shigella
	24	352	23.6	292504	1	AE016981 Shigella
	25	350.4	23.5	10206	1	AE015130 Shigella
	26	350.4	23.5	11626	1	AE015132 Shigella
	27	350.4	23.5	13143	1	AE015196 Shigella
	28	350.4	23.5	290319	1	AE016982 Shigella
	29	349.4	23.5	11320	1	AE010762 Methanosa
	30	348.8	23.4	17986	1	AF294823 Shigella
	31	346.8	23.3	321950	1	BX294138 Pirellula
	32	341.2	22.9	290582	1	AE016984 Shigella
	33	334.8	22.5	57655	1	AY316747_3
	34	334.6	22.5	302150	1	AP001510 Bacillus
	35	333.6	22.4	3289	1	AF486647 Rhizobium
	36	332.8	22.4	9872	1	ECO55214 Escherich
	37	331.8	22.3	10578	1	AE007314 Rhizobium
	38	326.2	21.9	4754	1	RL438039 Rhizobium
	39	323.2	21.7	340857	1	AP003010 Mesorhizo
	40	321.4	21.6	2223	6	CQ760129 Sequence
	41	318.4	21.4	297850	1	AP006577 Gloebact
	42	314.8	21.1	198050	1	AL646061 Ralstonia
	43	310.4	20.8	348971	1	BX572594 Rhodopneu
	44	309.8	20.8	298900	1	AP005937 Bradyrhiz
	45	309.2	20.8	2904	1	SRPRES Y14317 Streptomyce

ALIGNMENTS

RESULT 1	BD205211	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).	1489 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD205211					
DEFINITION	BD205211					
ACCESSION	BD205211					
VERSION	BD205211.1	GI:33014981				
KEYWORDS	JP 2002512813-A/1.					
SOURCE	unidentified					
ORGANISM	unclassified.					
REFERENCE	1 (bases 1 to 1489)					
AUTHORS	Frechon,D.T.M., Laure,F.C. and Thierry,D.					
TITLE	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)					
JOURNAL	Patent: JP 2002512813-A 1 08-MAY-2002;					
COMMENT	BIORAD PASTEUR					
	OS Unidentified					
	PN JP 2002512813-A/1					
	PD 08-MAY-2002					
	PF 27-APR-1999 JP 2000546051					
	PR 28-APR-1998 FR 98/05329					
	PI DOMINIQUE THERESE MARIE FRECHON,FRANCOISE CLAUDE LAURE, PI					
	PC CL2N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC					
	Strandedness: Double;					
	CC Topology: Linear;					
	CC Nucleotide sequence for detecting enterohemorrhagic CC					
	Escherichia coli					
	CC (EHEC).					
	PH Key					
	FT source					
	FT					
FEATURES	source					
	Location/Qualifiers					
	1..1489					
	/organism="unidentified"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:32644"					
ORIGIN						
	Query Match					
	Best Local Similarity					
	100.0%;					
	Score 1489;					
	DB 6;					
	Length 1489;					
	Pred. No. 0;					

Matches 1489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CTGCAGTCCGGAGATGAAGACCACTGTGTGTATACCCATACAGCTGTGTCCCGCAGGCCA	60
Db	1	CTGCAGTCCGGAGATGAAGACCACTGTGTGTATACCCATACAGCTGTGTCCCGCAGGCCA	60
QY	61	TGATTTTGTACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCCA	120
Db	61	TGATTTTGTACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCCA	120
QY	121	GCAGTTAAGCCAGCGGTTCCTGCTGAAATAATGTCCGCAAAACGGGGAAGCATCAGAA	180
Db	121	GCAGTTAAGCCAGCGGTTCCTGCTGAAATAATGTCCGCAAAACGGGGAAGCATCAGAA	180
QY	181	GGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC	240
Db	181	GGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC	240
QY	241	GGCGCTGATACCGGCAAGATGTGCGAATCTCCGCTCCGTGCAGCGGGCTATTTTCAGG	300
Db	241	GGCGCTGATACCGGCAAGATGTGCGAATCTCCGCTCCGTGCAGCGGGCTATTTTCAGG	300
QY	301	ATACCTTCTGTCAACACGTACAAACCAAGAACAGCAGCTTTTGTCTGACATCCAC	360
Db	301	ATACCTTCTGTCAACACGTACAAACCAAGAACAGCAGCTTTTGTCTGACATCCAC	360
QY	361	AAAGAAGGGAAATTCAGGTCTCGCAGCACTCAACGGCATCTGTCAGTTCCGGCTTGGA	420
Db	361	AAAGAAGGGAAATTCAGGTCTCGCAGCACTCAACGGCATCTGTCAGTTCCGGCTTGGA	420
QY	421	CCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTTGTCATCA	480
Db	421	CCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTTGTCATCA	480
QY	481	ATAAGTTGACTTTTGTATACCAATCGCAATTTCCCTTAATCCGAGCTATTCGTATGA	540
Db	481	ATAAGTTGACTTTTGTATACCAATCGCAATTTCCCTTAATCCGAGCTATTCGTATGA	540
QY	541	TAAAAAAACTCTTCCTGTTCTGATCTTCTGGCGCTATCGGGAGCTTTTCTACCGCTG	600
Db	541	TAAAAAAACTCTTCCTGTTCTGATCTTCTGGCGCTATCGGGAGCTTTTCTACCGCTG	600
QY	601	TAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAATC	660
Db	601	TAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTAATC	660
QY	661	CTCTGAGATTACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGATTATGCCACCA	720
Db	661	CTCTGAGATTACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGATTATGCCACCA	720
QY	721	GATTTCACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTT	780
Db	721	GATTTCACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTT	780
QY	781	CCAGGATTGGTCCCTCCGAGTATGGTCAATATGCTCTCTTTCTTTATTCGTATGCTT	840
Db	781	CCAGGATTGGTCCCTCCGAGTATGGTCAATATGCTCTCTTTCTTTATTCGTATGCTT	840
QY	841	GGCAGGTGCGGAAACATACAGGACATATATGCGGGAGGCGCAGTGGTGGTCAGC	900
Db	841	GGCAGGTGCGGAAACATACAGGACATATATGCGCGGGAGGCGCAGTGGTGGTCAGC	900
QY	901	AACGTTTTGAACCGCTGAACAGCTGGCGGATATACGTTAATCTGGATAAAGCCGTCGAT	960
Db	901	AACGTTTTGAACCGCTGAACAGCTGGCGGATATACGTTAATCTGGATAAAGCCGTCGAT	960
QY	961	TGCTGTGCCAGTCAAGAAAAATACCGCTCCAGTATTTCTCGGGAGACCTGATGTCC	1020
Db	961	TGCTGTGCCAGTCAAGAAAAATACCGCTCCAGTATTTCTCGGGAGACCTGATGTCC	1020
QY	1021	TGACTGTGTAATGTTGCCCTTGAATCCATGGGATTTAAAAACGCTGGGATTTCTGCGGAA	1080
Db	1021	TGACTGTGTAATGTTGCCCTTGAATCCATGGGATTTAAAAACGCTGGGATTTCTGCGGAA	1080

QY	1081	GAGAGATGACTGGGAGTCCGACCTGGTATACCTGGGGCTTGACAAAGCCTCTTTCAG	1140
Db	1081	GAGAGATGACTGGGAGTCCGACCTGGTATACCTGGGGCTTGACAAAGCCTCTTTCAG	1140
QY	1141	ATAACCGGATAAAAACGGAAACTTTCAGAACTCTTTCGGCCACGAGATGGGACTTA	1200
Db	1141	ATAACCGGATAAAAACGGAAACTTTCAGAACTCTTTCGGCCACGAGATGGGACTTA	1200
QY	1201	TTTATGTCAATCTGTAAGCCCCGGTGGAAACCCAGATCCTCTGGCTTCCGCCAAAGATA	1260
Db	1201	TTTATGTCAATCTGTAAGCCCCGGTGGAAACCCAGATCCTCTGGCTTCCGCCAAAGATA	1260
QY	1261	TCAGGGAAGCTTTTTCACGTATGGCCATGATGATGAGAGACTGTGGCCCTCATGCGCG	1320
Db	1261	TCAGGGAAGCTTTTTCACGTATGGCCATGATGATGAGAGACTGTGGCCCTCATGCGCG	1320
QY	1321	GAGGGATACATTTGGTAAAGCACATGGTGCAGCGTCTCTGAAAATGTTATTTGGCGCAG	1380
Db	1321	GAGGGATACATTTGGTAAAGCACATGGTGCAGCGTCTCTGAAAATGTTATTTGGCGCAG	1380
QY	1381	GGCTGATGGTGCACCTGTGGAGGACGAGGACTGGGATGGAATAATAATGTGTTACAG	1440
Db	1381	GGCTGATGGTGCACCTGTGGAGGACGAGGACTGGGATGGAATAATAATGTGTTACAG	1440
QY	1441	GAAACGGCAATATACCATCACCAGTGGCTCGAAGGAGCCTGCTCGAC	1489
Db	1441	GAAACGGCAATATACCATCACCAGTGGCTCGAAGGAGCCTGCTCGAC	1489
RESULT 2			
AX011297	LOCUS	AX011297	1489 bp DNA linear PAT 06-SEP-2000
AX011297	DEFINITION	Sequence 1 from Patent WO9955908.	
AX011297	ACCESSION	AX011297	
AX011297.1	VERSION	GI:9997847	
KEYWORDS	SOURCE	Escherichia coli	
ORGANISM		Escherichia coli	
REFERENCE		1	Thierry, D., Frechon, D.T. and Laure, F.C. Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)
AUTHORS			Patent: WO 9955908-A 1 04-NOV-1999;
TITLE			THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
JOURNAL			
FEATURES	source	1. 1489	Location/Qualifiers
		/organism="Escherichia coli"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:562"	
ORIGIN			
Query Match 100.0%; Score 1489; DB 6; Length 1489;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CTGCAGTCCGGAGATGAAGACCACTGTGTGTATACCCATCAGCGTGTCCCGCAGGCCA	60
Db	1	CTGCAGTCCGGAGATGAAGACCACTGTGTGTATACCCATCAGCGTGTCCCGCAGGCCA	60
QY	61	TGATTTTGTACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCCA	120
Db	61	TGATTTTGTACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCCA	120
QY	121	GCCAGTTAAGCCAGCGGTTCCTGCTGAAATAATGTCCGCAAAACGGGGAAGCATCAGAA	180
Db	121	GCCAGTTAAGCCAGCGGTTCCTGCTGAAATAATGTCCGCAAAACGGGGAAGCATCAGAA	180
QY	181	GGCGGGGGAACTCCGTCCGGCCAGTGAACCGTGCCACACTCCGGGCAGTACATGCCGC	240

Db 181 GGGCGGGGAACCTCGTCCGGCCAGTGAACCGTGCACACTCCGGCGAGTACATGCCGCC 240
Qy 241 GGGCGTGTATCGGCAAGATGTCGCAAACTCCCGCTCCGTGCGAGCGGCTATTTTCAGG 300
Db 241 GGGCGTGTATCGGCAAGATGTCGCAAACTCCCGCTCCGTGCGAGCGGCTATTTTCAGG 300
Qy 301 ATACCTTCGTTCATCAACACAGTACAAACACAGAACACAGCTTTTGTTCCTGACATCCAC 360
Db 301 ATACCTTCGTTCATCAACACAGTACAAACACAGAACACAGCTTTTGTTCCTGACATCCAC 360
Qy 361 AAAGAAGGAATATTCAGGCTGCGCAGCACTCAACGGCATCGTCAGTTGGGGCTTGGAA 420
Db 361 AAAGAAGGAATATTCAGGCTGCGCAGCACTCAACGGCATCGTCAGTTGGGGCTTGGAA 420
Qy 421 CCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTTGTCATCA 480
Db 421 CCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTTGTCATCA 480
Qy 481 ATAAAGTTGACTTTTGTATACAAATGGAATTTCCCTTAATCCGGAGCTATTCGTATGA 540
Db 481 ATAAAGTTGACTTTTGTATACAAATGGAATTTCCCTTAATCCGGAGCTATTCGTATGA 540
Qy 541 TAAAAAATCTTCTCTGTTCTGATTTCTTCTGGGCTATCCGGGAGCTTTTCTACCGCTG 600
Db 541 TAAAAAATCTTCTCTGTTCTGATTTCTTCTGGGCTATCCGGGAGCTTTTCTACCGCTG 600
Qy 601 TAGCGCTGTATAAAGAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTC 660
Db 601 TAGCGCTGTATAAAGAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTC 660
Qy 661 CTCTGAGATTACAGCCCTCAATCAATCCCTGGGGCTGATTTGATTTATGCCACCA 720
Db 661 CTCTGAGATTACAGCCCTCAATCAATCCCTGGGGCTGATTTGATTTATGCCACCA 720
Qy 721 GATTTCACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGCAAACTT 780
Db 721 GATTTCACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTGCTGCAAACTT 780
Qy 781 CCCAGGATGTCGCCCTGCGGATTTATGTCATATATGTCCTTTCTTTATTCGTATGGCTT 840
Db 781 CCCAGGATGTCGCCCTGCGGATTTATGTCATATATGTCCTTTCTTTATTCGTATGGCTT 840
Qy 841 GGACAGTGGCGGAACATACAGGACATATGATGCGCGGGAGGCCAGTGGTGGTCCAGC 900
Db 841 GGACAGTGGCGGAACATACAGGACATATGATGCGCGGGAGGCCAGTGGTGGTCCAGC 900
Qy 901 AACGTTTTGAACCGCTGAACAGCTGGCGGATAACGTTAACTCGATAAAGCCGCTCGAT 960
Db 901 AACGTTTTGAACCGCTGAACAGCTGGCGGATAACGTTAACTCGATAAAGCCGCTCGAT 960
Qy 961 TGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTCGGGAGACCTGATGTC 1020
Db 961 TGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTCGGGAGACCTGATGTC 1020
Qy 1021 TGACTGTATGTTGCCCTTGAATCCATGGATTTAAACGCTGGATTTGCTGCGGAA 1080
Db 1021 TGACTGTATGTTGCCCTTGAATCCATGGATTTAAACGCTGGATTTGCTGCGGAA 1080
Qy 1081 GAGAAGATGACTGGAGTCCGACCTGGTATATCTGGGGCTGACAAAGGCTCTTTCGAG 1140
Db 1081 GAGAAGATGACTGGAGTCCGACCTGGTATATCTGGGGCTGACAAAGGCTCTTTCGAG 1140
Qy 1141 ATAAACGGGATAAAGACGGGAACTTCAGAACTCTTTCGCCCAACGAGATGGGACTTA 1200
Db 1141 ATAAACGGGATAAAGACGGGAACTTCAGAACTCTTTCGCCCAACGAGATGGGACTTA 1200
Qy 1201 TTTATGTCATCTCGAGGCCCCGGTGGAAAAACAGATCTCTCTGCTTCCCGGAAAGATA 1260
Db 1201 TTTATGTCATCTCGAGGCCCCGGTGGAAAAACAGATCTCTCTGCTTCCCGGAAAGATA 1260
Qy 1261 TCAGGGAAGCTTTTTCAGCTATGCGCATGATGATGAGGAGCTGTGGCCCTGATCGGG 1320
Db 1261 TCAGGGAAGCTTTTTCAGCTATGCGCATGATGATGAGGAGCTGTGGCCCTGATCGGG 1320

Qy 1321 GAGGCGCATACATTTGGTTAAAGCACATGCTGAGCGTCTCTCGAAAAATGTTATTGGCGCAG 1380
Db 1321 GAGGCGCATACATTTGGTTAAAGCACATGCTGAGCGTCTCTCGAAAAATGTTATTGGCGCAG 1380
Qy 1381 GGCCTGATGGTGCACCTGCTGAGGAGCAGGACTGGGATGGAATAAATCTGCTACAG 1440
Db 1381 GGCCTGATGGTGCACCTGCTGAGGAGCAGGACTGGGATGGAATAAATCTGCTACAG 1440
Qy 1441 GAAACGGCAATATACCATCACTACCGTGGCTGGAGGAGCCTGGTTCGAC 1489
Db 1441 GAAACGGCAATATACCATCACTACCGTGGCTGGAGGAGCCTGGTTCGAC 1489

RESULT 3
AF074613 92077 bp DNA circular BCT 04-NOV-1998
LOCUS Escherichia coli O157:H7 plasmid pO157, complete sequence.
DEFINITION AF074613
ACCESSION AF074613
VERSION AF074613.1 GI:3822114
SOURCE
ORGANISM Escherichia coli O157:H7
Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 92077)
AUTHORS Burland, V., Shao, Y., Perna, N.T., Plunkett, G., Sofia, H.J. and Blattner, F.R.
TITLE The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7
JOURNAL Nucleic Acids Res. 26 (18), 4196-4204 (1998)
MEDLINE 98391744
PUBMED 9722640
REFERENCE 2 (bases 1 to 92077)
AUTHORS Burland, V., Shao, Y., Perna, N.T., Plunkett, G. III, Sofia, H.J. and Blattner, F.R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. 92077
/organism="Escherichia coli O157:H7"
/mol_type="genomic DNA"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:83334"
/lab_host="Escherichia coli C600"
/plasmid="pO157"
1. 561
/gene="f1n0"
/note="synonym: L7001"
1. 561
/gene="f1n0"
/note="97 pct identical amino acid sequence and equal length to F101_ECOLI SW: P22707"
/codon_start=1
/transl_table=11
/product="fertility inhibition protein (conjugal transfer repressor)"
/protein_id="AAC70069.1"
/db_xref="GI:3822115"
/translation="MAEQKRPVLTKRTEGETPVRSRKTIIVTTPPKVKQKLA EKAREALAAKQAQALSIYINLTODEAVNTLKPWPLGDFDGTPLLAGCIRD VLLEDVQRNIPLSHKKRLRAIKATRSYSLCAMKACAKRYDTEGYVTEHISQEEA YAAARDKIRRRQNRIRKAEQAVLDEK"
653..681
/note="predicted sigma 70 promoter; score of 52%"
698..949
/gene="L7002"
698..949
/gene="L7002"
/note="54 pct identical (1 gap) to 23 residues of an approx. 269 aa protein EXOK_RHIME SW: P33693"

gene
CDS
promoter
gene
CDS

```
promoter
gene
CDS

/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC70070.1"
/db_xref="GI:3822116"
/translations="MSETVHGTVRSQVTSVPAGPLFWKSVDAWMKRQKHGDLPLV
HPGLTSSLPKGLNTATGAABEGGNEKSLHSRDSEG"
1025..1052
/notes="predicted sigma 70 promoter; score of 62%"
1151..1612
/gene="L7003"
1151..1612
/gene="L7003"
/notes="98 pct identical and equal length to YFI3_ECOLI SW:
Q93342"
/codon_start=1
/transl_table=11
/product="hypothetical protein 15.6 kDa protein in finO 3'
region precursor"
/protein_id="AAC70071.1"
/db_xref="GI:3822117"
/translations="MRXIPLVLFISFVPLCADIHGRVVRVLDGDTIEVMSRKA
IRLVNIDAPKKQDYGKSDMMKSLVAGKTVTVYFQDRYGRMLQGVYAPDGMV
QFVRAAGAAVYEQYNTDPLVPLVQLNEARQQKGLMSDADVPFPWIWRHK"
1390..1419
/notes="predicted sigma 70 promoter; score of 56%"
1658..1867
/gene="L7004"
1658..1867
/gene="L7004"
/notes="55 pct identical (0 gaps) to 66 residues of an
approx. 72 aa protein HHA_ECOLI SW: P23870"
/codon_start=1
/transl_table=11
/product="putative hemolysin expression modulating
protein"
/protein_id="AAC70072.1"
/db_xref="GI:3822118"
/translations="WERTKQEWLYQLRCSVNTLEKIIHKNRDSLTSRSPNSA
DHRLAELITGLKYDRIPKEIKYVR"
1899..1927
/notes="predicted sigma 70 promoter; score of 56%"
1905..2243
/gene="L7005"
1905..2243
/gene="L7005"
/notes="98 pct identical (0 gaps) to 57 residues of an
approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF
4, TRSW: Q52340"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC70073.1"
/db_xref="GI:3822119"
/translations="WKLIIFTLIIIVLIIAALLIRILRSVNHQSPLLQLHAAGIRTD
AERILSGEYDASHLRPERRRDILLEVLVLRQAGIPLLRSHDARKLLQMTGEWLNTTG
AAQSPGHR"
misc_feature
/notes="initiation site; Escherichia coli plasmid R100 ssIB
gene; 99 pct identical to ssIB locus R100SIB accession
D90185"
2416..2421
/notes="100 pct identical (0 gaps) to the -35 region at
23..28 locus ECNR1REP accession X02302"
2439..2445
/notes="100 pct identical (0 gaps) to the -10 region at
46..52 locus ECNR1REP accession X02302"
2473..2478
/notes="100 pct identical (0 gaps) to RBS at 80..85 locus
ECNR1REP accession X02302"
2483..2737
/gene="cpb2"
/notes="synonym: L7006"

promoter
gene
CDS

/codon_start=1
/transl_table=11
/product="CpB2 protein (RepA2 protein)"
/protein_id="AAC70074.1"
/db_xref="GI:3822120"
/translations="MSQTENAVTSSSGAKAYBKGNPLSDAEKQLSVARKSFKEV
KVLEPKYKAMLMQMCHDGLTQAEVLTAIKSEAQRKV"
2754..2782
/notes="predicted sigma 70 promoter; score of 69%"
2973..3047
/gene="L7007"
2973..3047
/gene="L7007"
/notes="91 pct identical (0 gaps) to 24 residues of an
approx. 24 aa protein uORF P30REPFIC, accession M16167,
translationally coupled to replication initiation protein"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70075.1"
/db_xref="GI:3822121"
/translations="MLGKVQDFFLCSELLRIVSAGWCD"
3032..3035
/notes="100 pct identical (0 gaps) to RBS at 641..644 locus
ECNR1REP accession X02302"
3040..3897
/gene="rep2"
/notes="synonym: L7008"
3040..3897
/gene="rep2"
/notes="96 pct identical amino acid sequence and equal
length to REP2_ECOLI SW: P03066"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/translations="MTDLQTYRQVKNPNVFTFREGAGTLKCEKMEKAVGFTSR
FDFAIHVHARSKGLRRMPVLRRAIDALLQGLCFHDFPLANRVQCSITFLAIECG
LATSAAGKLSITRATRALTFLSELGITVQTEYDITGCVITDITFTALFAALDV
SEDAVAARSRVEMENLRKKOGLDTLGMDELIAKAWRFVRFRPSYQTELKSGIK
RARARDAGRQRQDITLVKRLQLTREISEGRFANREAVKREVRVYKERMILSEN
YSRLATASP"
3847..4095
/standard_name="CIS"
/notes="88 pct identical to locus ECCIS accession X12587,
required for cis-activation of oriR by the replication
initiation protein"
4072..4080
/notes="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1REP at (1682..1690) accession X02302"
4084..4232
/standard_name="oriP"
/notes="89 pct identical to oriR (1094..1242); minimum
segment for replication of E. coli IncFII plasmid NR1
ECREPA1 X12776"
4169..4197
/notes="predicted sigma 70 promoter; score of 56%"
4259..4453
/gene="L7009"
4259..4453
/gene="L7009"
/notes="95 pct identical to (0 gaps) 64 residues of a 128
aa protein RBP4 locus ECRSC1 accession V00351"
/codon_start=1

Query Match          99.2%; Score 1476.4; DB 1; Length 92077;
Best Local Similarity 99.9%; Pred. No. 0;
```

Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	1	CTGCAGTCCGAGATGAA-AGCACCACCTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC	59
Db	6836	CTGCAGTCCGAGATGAAACAGCAACCTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC	6895
Qy	60	ATGATTTTGTGCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC	119
Db	6896	ATGATTTTGTGCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC	6955
Qy	120	AGCCAGTTAAGCCAGCGGTTTCCCTGCTGTAATAATGTGCGCAAAACCGGGGAAGCATCAGA	179
Db	6956	AGCCAGTTAAGCCAGCGGTTTCCCTGCTGTAATAATGTGCGCAAAACCGGGGAAGCATCAGA	7015
Qy	180	AGGCGGGGGAATCCCTCGGCGAGTGAACCGTCCGACACTCCGGGAGTACATGCGC	239
Db	7016	AGGCGGGGGAATCCCTCGGCGAGTGAACCGTCCGACACTCCGGGAGTACATGCGC	7075
Qy	240	CGGCGCTGATACCGGCAAGATGTGCAAACTCCCGTCCGTCAGCGGCTATTTTCAG	299
Db	7076	CGGCGCTGATACCGGCAAGATGTGCAAACTCCCGTCCGTCAGCGGCTATTTTCAG	7135
Qy	300	GATACCTTCTGTCATCAACAGCTAACAAACAGAAAGACAGCTTTTGTGTTCTGACATCCA	359
Db	7136	GATACCTTCTGTCATCAACAGCTAACAAACAGAAAGACAGCTTTTGTGTTCTGACATCCA	7195
Qy	360	CABAAGAGGAATATTGAGTCTGCGCAGCACTCAACGGGCATCGTCAAGTTCGGCTTGA	419
Db	7196	CABAAGAGGAATATTGAGTCTGCGCAGCACTCAACGGGCATCGTCAAGTTCGGCTTGA	7255
Qy	420	ACCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTTGCATC	479
Db	7256	ACCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTTGCATC	7315
Qy	480	AATAAGTTGACTTTTGTATACAACTACAAACAGAAAGACAGCTTTTGTGTTCTGACATCCA	539
Db	7316	AATAAGTTGACTTTTGTATACAACTACAAACAGAAAGACAGCTTTTGTGTTCTGACATCCA	7375
Qy	540	ATAAAAAAACTCTCTGTTCTGATTTCTGCGCTATCTGCGGCTATCGGGAGCTTTTACCGCT	599
Db	7376	ATAAAAAAACTCTCTGTTCTGATTTCTTCTGCGCTATCGGGAGCTTTTCTACCGCT	7435
Qy	600	GTAGCCCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT	659
Db	7436	GTAGCCCTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT	7495
Qy	660	CTCTGAGATTACACAGCCCTGAAATCAAAATCCCTGGGGGCTGATTTTGAATATGCCACC	719
Db	7496	CTCTGAGATTACACAGCCCTGAAATCAAAATCCCTGGGGGCTGATTTTGAATATGCCACC	7555
Qy	720	AGATTTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACT	779
Db	7556	AGATTTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAAAGATTTGCTGACAACT	7615
Qy	780	TCCAGGATTTGGTCTCGGGATTTATGTCATTATGCTCTCTTTTATTTATGCTATGGCT	839
Db	7616	TCCAGGATTTGGTCTCGGGATTTATGTCATTATGCTCTCTTTTATTTATGCTATGGCT	7675
Qy	840	TGCAACGCTCCGGAACATACAGACATATGATGGCGGGGAGCGCCAGTGGTGTGCTAG	899
Db	7676	TGCAACGCTCCGGAACATACAGACATATGATGGCGGGGAGCGCCAGTGGTGTGCTAG	7735
Qy	900	CAACGTTTGAACCGCTGAAACAGTGGCCGGATTAACGTTTATCTGATTAAGCCCGTCCA	959
Db	7736	CAACGTTTGAACCGCTGAAACAGTGGCCGGATTAACGTTTATCTGATTAAGCCCGTCCA	7795
Qy	960	TTCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTC	1019
Db	7796	TTCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTC	7855
Qy	1020	CTGACTGGTAAATGTGCTTGAATCCATGGGATTTTAAAAACGCTGGGATTTGCTGGCGGA	1079
Db	7856	CTGACTGGTAAATGTGCTTGAATCCATGGGATTTTAAAAACGCTGGGATTTGCTGGCGGA	7915

Qy	1080	AGAGAAGATGACTGGGAGTCCGACCTGCTATATCTGGGGGCTTGACAAACAAGCCTCTTTGCA	1139
Db	7916	AGAGAAGATGACTGGGAGTCCGACCTGCTATATCTGGGGGCTTGACAAACAAGCCTCTTTGCA	7975
Qy	1140	GATAACCGGATATAAAACGGGAACTTTCAGAAAACCTCTTGCAGCAGCGAGATGGGACTT	1199
Db	7976	GATAACCGGATATAAAACGGGAACTTTCAGAAAACCTCTTGCAGCAGCGAGATGGGACTT	8035
Qy	1200	ATTATGTCAATCTCTGAAGCCCGGTTGAAAAACAGATCTCTTGGCTTCCGCGAAAGAT	1259
Db	8036	ATTATGTCAATCTCTGAAGCCCGGTTGAAAAACAGATCTCTTGGCTTCCGCGAAAGAT	8095
Qy	1260	ATCAGGGAAGCTTTTTCACGTATGGCCATGATGATGAGAGACTGTGGCCCTGATCGCG	1319
Db	8096	ATCAGGGAAGCTTTTTCACGTATGGCCATGATGATGAGAGACTGTGGCCCTGATCGCG	8155
Qy	1320	GGAGGGCATACATTTGGTAAAGCACATGTCGACGCTCTCTGAAAAATGTATTGGCGCA	1379
Db	8156	GGAGGGCATACATTTGGTAAAGCACATGTCGACGCTCTCTGAAAAATGTATTGGCGCA	8215
Qy	1380	GGGCTGTGATGCTCACCTGTGGAGGACAGGAGCTGGGATGGAATAAATGTGTGTACA	1439
Db	8216	GGGCTGTGATGCTCACCTGTGGAGGACAGGAGCTGGGATGGAATAAATGTGTGTACA	8275
Qy	1440	GGAAACGGCAAAATATACCATCACAGTGGCTGGAAGGAGCCTGGTGCAC	1489
Db	8276	GGAAACGGCAAAATATACCATCACAGTGGCTGGAAGGAGCCTGGTGCAC	8325
RESULT 4			
AX191727	LOCUS	AX191727	92077 bp DNA linear PAT 15-AUG-2001
DEFINITION	Sequence 9 from Patent WO0149775.		
ACCESSION	AX191727		
VERSION	AX191727.1	GI:15209896	
KEYWORDS	Escherichia coli		
SOURCE	Escherichia coli		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.		
REFERENCE	1	Iversen, P.L.	
AUTHORS	Antisense antibacterial cell division composition and method		
TITLE	Patent: WO 0149775-A 9 12-JUL-2001;		
JOURNAL	Avi Biopharma, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..92077	/organism="Escherichia coli"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:562"	
ORIGIN			
Query Match	99.2%;	Score 1476.4;	DB 6; Length 92077;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 1488;	Conservative 0;	Mismatches 1;	Indels 1; Gaps 1;
Qy	1	CTGCAGTCCGAGATGAA-AGCACCACCTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC	59
Db	6836	CTGCAGTCCGAGATGAAACAGCAACCTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC	6895
Qy	60	ATGATTTTGTGCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC	119
Db	6896	ATGATTTTGTGCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC	6955
Qy	120	AGCCAGTTAAGCCAGCGGTTTCCCTGCTGTAATAATGTGCGCAAAACCGGGGAAGCATCAGA	179
Db	6956	AGCCAGTTAAGCCAGCGGTTTCCCTGCTGTAATAATGTGCGCAAAACCGGGGAAGCATCAGA	7015
Qy	180	AGGCGGGGGAATCCCGTCCGGCAGTGAACCGTGCACACTCCGGGAGTACATGCGCG	239
Db	7016	AGGCGGGGGAATCCCGTCCGGCAGTGAACCGTGCACACTCCGGGAGTACATGCGCG	7075

QY 240 CGCGCGCTGATACCGGCAAGAAATGGTCGCAAACTCCCGCTCCGTCGAGCGGCTATTTCAG 299
DB |||||
QY 7076 CGCGCGCTGATACCGGCAAGAAATGGTCGCAAACTCCCGCTCCGTCGAGCGGCTATTTCAG 7135
DB |||||
QY 300 GATACCCCTTCGTCATCAACACGTCACAAACAGAGACCAAGCTTTTGTTCGACATCCA 359
DB |||||
QY 7136 GATACCCCTTCGTCATCAACACGTCACAAACAGAGACCAAGCTTTTGTTCGACATCCA 7195
DB |||||
QY 360 CAAAGAGGGAATATTTCAGTCTCGGAGACGCTCAACGGGATCGTCAGTTGCGGCTTGA 419
DB |||||
QY 7196 CAAAGAGGGAATATTTCAGTCTCGGAGACGCTCAACGGGATCGTCAGTTGCGGCTTGA 7255
DB |||||
QY 420 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTCTGTTCATC 479
DB |||||
QY 7256 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTCTGTTCATC 7315
DB |||||
QY 480 AATAAAGTTGACTTTTGTATACAACTCGAAATTTCCCTTAATCCGAGCTATTTCGTATG 539
DB |||||
QY 7316 AATAAAGTTGACTTTTGTATACAACTCGAAATTTCCCTTAATCCGAGCTATTTCGTATG 7375
DB |||||
QY 540 AATAAAGTTGACTTTTGTATACAACTCGAAATTTCCCTTAATCCGAGCTATTTCGTATG 599
DB |||||
QY 7376 AATAAAGTTGACTTTTGTATACAACTCGAAATTTCCCTTAATCCGAGCTATTTCGTATG 7435
DB |||||
QY 600 GTAGCCGCTGATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT 659
DB |||||
QY 7436 GTAGCCGCTGATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACT 7495
DB |||||
QY 660 CCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTTGTATTATGCCACC 719
DB |||||
QY 7496 CCTCTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTTGTATTATGCCACC 7555
DB |||||
QY 720 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAGATATCAAGATTTTGTGACAACT 779
DB |||||
QY 7556 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAGATATCAAGATTTTGTGACAACT 7615
DB |||||
QY 780 TCCAGGATTTGTCCTGCGGATTTATGCTTATGCTCTCTTCTTTATTCGTATGGCT 839
DB |||||
QY 7616 TCCAGGATTTGTCCTGCGGATTTATGCTTATGCTCTCTTCTTTATTCGTATGGCT 7675
DB |||||
QY 840 TGGCAGCGTCCGGAACATACAGGACATATGATGGCCGGGAGCGCAGTGGTGGTFCAG 899
DB |||||
QY 7676 TGGCAGCGTCCGGAACATACAGGACATATGATGGCCGGGAGCGCAGTGGTGGTFCAG 7735
DB |||||
QY 900 CAAGCTTTTGAACCGCTGAAACAGCTGGCCGGAATACGTTTAACTCTGGATAAAGCCGCTCGA 959
DB |||||
QY 7736 CAAGCTTTTGAACCGCTGAAACAGCTGGCCGGAATACGTTTAACTCTGGATAAAGCCGCTCGA 7795
DB |||||
QY 960 TTGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTTCTTGGGAGACCTGATGGTC 1019
DB |||||
QY 7796 TTGCTGTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTTCTTGGGAGACCTGATGGTC 7855
DB |||||
QY 1020 CTGACTGGTAATGTTGCCCTTGGATCCATCGGATTTTAAACGCTGGGATTTGCTGGCGGA 1079
DB |||||
QY 7856 CTGACTGGTAATGTTGCCCTTGGATCCATCGGATTTTAAACGCTGGGATTTGCTGGCGGA 7915
DB |||||
QY 1080 AGAGAAATGATGCTGGAGTGGACCTGGTATATCTGGGGGCTGACAAAGACCTCTTTCGA 1139
DB |||||
QY 7916 AGAGAAATGATGCTGGAGTGGACCTGGTATATCTGGGGGCTGACAAAGACCTCTTTCGA 7975
DB |||||
QY 1140 GATACCCGGGATAAAGCGGAAACTTTCAGAAACCTCTTGGCGGACCGAGATGGGACTT 1199
DB |||||
QY 7976 GATACCCGGGATAAAGCGGAAACTTTCAGAAACCTCTTGGCGGACCGAGATGGGACTT 8035
DB |||||
QY 1200 ATTTATGTCATCTCAAGGCCCGGCTGGAAACACAGATCTCTGGCTTCGCGGAAAGAT 1259
DB |||||
QY 8036 ATTTATGTCATCTCAAGGCCCGGCTGGAAACACAGATCTCTGGCTTCGCGGAAAGAT 8095
DB |||||
QY 1260 ATCAGGAAGCTTTTTCAGTATGGCATGATGATGAGGAGCTGTGGCCCTGATCGCG 1319
DB |||||
QY 8096 ATCAGGAAGCTTTTTCAGTATGGCATGATGATGAGGAGCTGTGGCCCTGATCGCG 8155
DB |||||
QY 1320 GGAGGGCATACATTTGGTAAAGCATGGTTCAGCGCTCTCTCGAAAAATGATTGGCGCA 1379

DB 8156 GGAGGGCATACATTTGGTAAAGCACATGGTCAGCGCTCTCTGAAAAATGATTGGCGCA 8215
QY 1380 GGGCCTGATGGTCACCTGTGGAGGAGCAGGACTCGGATGGAATAAATGTTGGTACA 1439
DB 8216 GGGCCTGATGGTCACCTGTGGAGGAGCAGGACTCGGATGGAATAAATGTTGGTACA 8275
QY 1440 GGAACGGCAATATATACCATCACCTGGCTTGGAAAGAGCCTGGTTCGAC 1489
DB 8276 GGAACGGCAATATATACCATCACCTGGCTTGGAAAGAGCCTGGTTCGAC 8325

RESULT 5
LOCUS AB011549 92721 bp DNA circular BCT 27-APR-1999
DEFINITION Escherichia coli plasmid p0157 DNA, complete sequence.
ACCESSION AB011549
VERSION AB011549.2 GI:4589740
KEYWORDS ToxR-regulated lipoprotein; tagA.
SOURCE Escherichia coli
ORGANISM Escherichia coli
REFERENCE 1 (sites)
AUTHORS Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K., Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K., Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S. and Shinagawa,H.
Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic Escherichia coli O157:H7 derived from Sakai outbreak
DNA Res. 5 (1), 1-9 (1998)
MEDLINE 98290540
PUBMED 9628576
REFERENCE 2 (bases 1 to 92721)
AUTHORS Makino,K.
TITLE Direct Submission
JOURNAL Submitted (24-PEB-1998) Kozo Makino, Research Institute for Microbial Diseases, Osaka University, Molecular Microbiology, Yamadaoka, 3-1, Suita, Osaka 562, Japan
EX-MAIL:makino@obkns01.biken.osaka-u.ac.jp, Tel:81-6-879-8318, Fax:81-6-879-8320
COMMENT On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES
Location/Qualifiers
1..92721
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:562"
/plasmid="p0157"
/note="RIMD 0509952 is a strain of enterohemorrhagic E. coli, EHEC O157:H7"
join(92527..92721,1..2502)
/gene="tagA"
join(92527..92721,1..2502)
/gene="tagA"
/codon_start=1
/transl_table=11
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4666293"
/translation="MNTKNRWRTPMKLVSLCTILAPLAIGVFSATAADNNSAIYFNTSQPNDLQGLAAEVKFAQSOILPAHPKGDOSPHTLSRLSLILAVRPVKADKTPVQVEARDNNKILGTLTPSPSLPTDIYHLDGVPPEGIDFTPHNGTKKLINTVAEYNKLSDSGSSIHSLTNNALVEIHTANGWRWDIYLPQGPDLGQWVRVFSAGYSTV FYGDRKVTLSVGNLTLFYKVGQWFRSGELENNRIYQAHIWSAELPAHWIYVGLNLV I KGNLSGRINDIKIGAPGELLHTIDIGMLTTPRDRFDPAKDEAREHYFQITPVSRI MIYNNYAPLHLKEVMTGLLMDPGNGWHSQWTRORIGKELVSHGIDNANYGLN STAGLGNHPYVVAQLAAHNSRGNVANGIQVHGGSGGGIVTLDSFTLGNFPEVHGHN YGLGVHDYDFKGSVHRSAENNNSTWGDGDKRFPINFPYPSQTNEKSLNNQCEPFDGHKFGFDNAGGSPFSAANRFTMYTPNSAIIORFENKAVFDSDRSSTGFSKKNADT QMEPEYHTIDRAEQITASVNELSESMAELMAEYAVVVKHMMNGNWNTRNIYIPTASA

DNRGSIITINHEAGVNSYLFINGDEKVVSYQYKKSFYSDGQFWKRDVVDVTEARKPE
QGVPTVTLVGYDPEGTLSSYIVPAMYGYAGFTYSDDSQNLSDNDQLOVDYTKEGOL
RFLRANRANTVWKNKHINPTESQPTQATLVCKNNKILDTKSLTPAPEGLTYTVNGQ
ALPARENKGCIVNSGKRYCLPVGQSGYSLPDWIVGQBYVDSGAKAKVLLSDWDN
LSYNRIGFGBVGNVNPDMKKYKAWNGQYLDSPKPSMRVYK"

gene
CDS

2589..3464
/gene="etpc"
2589..3464
/gene="etpc"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31758.1"
/db_xref="GI:3336999"

/translation="MLPFLSRGRLFIKIDVLKMLTPNRLLCVILLIAGYQLVSVI
HFWLTAQASVGLSRVSAPEVATVQGTQERFVTLFGRASPLSSGSAQETWPSLUS
DOLLSGDLVDVGLYSVAHSAVIAFAHNRQFSLVSGEKVSYDIATISAFSDHVS
INYOQTVSLPRYDNTKKNAYDNMNLTVGDVITQDNFRVESVFDIMSFAVTNNVT
LSGYRLIPGKHSLFYNAGLHNDLAVSVNGSELDRDTRQAQIMKQLPELKEITVE
RDGQLYDAFIAVGEN"

gene
CDS

3675..5432
/gene="etpb"
3675..5432
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31759.1"
/db_xref="GI:3337000"

/translation="MLNEEQYQFSLVDVYGAIVDMHNGILKVVRSKDAKTSVAP
VASDVPQTGDEVTVRVVPSNVAARDLAPLLROLNDNAGASVVHVPSNVLMTGR
AAVMKRLMEIVERDKVGNRSVATVPLTYASATDVARLVLTETKTUKTALPAMWTAK
LVADERTNSVLVSGEPISSQRIISIIKQDQEDVQGNKVIYIKYAKADOLVEVLIG
ISSIENDSKSPSTEALRGVITKSHQETNALILTAPDVIRLNVISOLDIRRPQ
LVBAIIAEGDADGLNLGIQVNNKHAGVAQFTSTGLPTITTMQTRONEILDSQNSA
LSMFGAAGYQGNWAMLLTALSTSKNDILATPSIVTIDNMEATFNVEGVPEVLSG
SOTFGDNIENFWKRTVGIKLRVKPOINEGDSVLEIEQEVSGVADTAVATTTDLGA
TNTFTVTNMLVNGGVVVGGLLDSIRGSESKVPLLGIDIPVLGHLFRAKSQSTAK
RNMLFRTPIIRDRGFRHASAEKIOSFQEQVQSRGKETTLTALTLNEEQURLSPDQD
DTAFRKVKAIAFPAQEM"

gene
CDS

5432..6937
/gene="etpb"
5432..6937
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"

/translation="MSRVQNVSESRPLLPFSRSTORILLRREQGNRVFCMEDTGA
SALLEVRVAREGPNLTVTVAEAEKQLVSSYQRDSDEARQMMABEIGNEMDFYTVAGE
LPDREDLDDANDDAPIIRLINAMLTAEIKKASDIHETVYERHQLVRFIDGVILREIL
RLHRNLASLITSRKVMARDIAEKRPVQDGRMVLIRIGRAVDVRVSTLPSNHGERIV
LRLLDKNSVLDLAALGMSQONRHIDALIRRHGIIIVTGPTGSKSTTLIYAALSLL
NPRDNTVVEDPVEYELDGI SQTVNPKVDMTFARSILRAILRQDPDVLVEGRDGE
TAQIAVQASLTGHLVLTHTNSAAGALSRLQDMGIPFPLLSLVAQLRVLTLIC
PRYQPCQVSTELAMDIDIPETITWQPCQHCSTGYHGTGRIHEILLIDIRTA
IYQGEGLGHTRLAGSYRLTRGDGRQKVLAGESSWEVVRVTSRLQEEB"

gene
CDS

6939..8162
/gene="etpb"
6939..8162
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"

/translation="MALFHYQASDIHGRKSGILEADSARHARQLLREQALIPVRLDE
KQVHKHSRLSILFRPRGSSAEALLTQLATLVAAASLPLEALDALLRQSKPKQ
RNLTAAVTRTKVLEGHSAAMGPFQFERYLCAMVAAGETSGRLDVLVLSRLADYTPQ
ROIENRRLQLLPCVLTIVAVGVIAILLTAVVPKVVEQFIHMKQTLPLSTRVLMGA
AEVSTQWGPMLLAAAGGLTAGRMILHOPSORLAFHLLRLPLVGVGRISGLNTARVA
RTLILNASAVPLLQAMHISGDVLSDNWARHQLATAELVREGVSLHQALEQTSILFPP

MRRHMIASGENSGELDSMLERAADNQDREFSTQMLALGLPEPLLVMGAGVVLFIIVL
AIIQPLLIQLNNMNM"

gene
CDS

8193..8627
/gene="etpb"
8193..8627
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31762.1"
/db_xref="GI:3337003"

/translation="MRKHQHGFTLLIIMVIVILGLASLVNPNLMGNKDKADROKV
MSDIALVLESTLDWYRLDNNRYPTTEQGLRALVSKPTVQPEPRNRYQDGYIRRLQDPW
GSDYQLLNPQYSIDIDIFSPGPGVPTNEDDIGNWTIGNAQP"

gene
CDS

8624..9178
/gene="etpb"
8624..9178
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337004"

/translation="MSQRGFTILEMMLILLMGTAGLVLMSPDPSAQNHQQOQERL
QAOLDYALDRSQDGLLGMIOVRDGMKFKVLORGTAESPPTLAEGGDIWQYVQW
QPRAMGKGLPDVRLQLYLRGLQWMSEHDDGAEFDILLPGCVTPFPFLLRQVG
EEAVVGLQVDENGLMTLFEGEVSL"

gene
CDS

9193..9540
/gene="etpi"
9193..9540
/gene="etpi"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31764.1"
/db_xref="GI:3337005"

/translation="MTLEVLVAVFALAGMALMQASTQQAAGIGRMEEKVLGWL
DNQVQLQKSTWPEWNGWGEKTSIFAGTEWYLRWGGPDSVQPRSLSEVVRKEET
SALVLSURKSVVRE"

gene
CDS

9537..10136
/gene="etpb"
9537..10136
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31765.1"
/db_xref="GI:3337006"

/translation="MSQQRVKGFTLLEMLLAVFAALSISAFQVLOGGIRAHLSRD
KVORLAEQLQISQMERDLTQMLPRHSGNEGGLLAPHLKSDDDMGISFTRNSWLN
AGMLPRPELQWGYRLRQKLERLSYFVHDHPSGSPDVRLMDGVHAFRLRFFVNGD
WQARWSTGLLPQAVEVTLVMDDFAEPLRLFLSVKETA"

gene
CDS

10133..11110
/gene="etpb"
10133..11110
/gene="etpb"
/codon_start=1

/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31766.1"
/db_xref="GI:3337007"

/translation="MKUREQVALLVLLILSLMVTIAAVIAERNGTFTLRTVAQLDO
LOAKWDGYTAETIAKQILORSRSPKTHLAQNWAQSERQFETRGDVRGQIVDAQ
CFNLNAINYGVVDLTISPYAARIFQQLINLQVELQARVTAALRWDIRDDEKPVRG
GABDEVYMGEPVPLAANQMDVSELRLIGLDARLYKLLPVYCLVSLSVNVN
TLDSQAPLLAALFLTKPDSLPTVELLOREPRTGWESVAFLDAPLKDIDTSAAMPV
LAVSNYFLVRLHVRSGEHLFSQOTLMQWREERFIQRQYGLTMRVP"

11473..12321

gene
Query Match
Best Local Similarity
Matches 1488, Conservative

99.2%; Score 1476.4; DB 1; Length 92721;
99.9%; Pred. No. 0;
0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 CTGCAGTCCGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
Db 76351 CTGCAGTCCGAGATGAACAGACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 76410
Qy 60 ATGATTTTGTTCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC 119
Db 76411 ATGATTTTGTTCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC 76470
Qy 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGTCCGCAAAACCGGGGAAGCATCAGA 179
Db 76471 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGTCCGCAAAACCGGGGAAGCATCAGA 76530
Qy 180 AGGCGGGGGAACCTCCGTCGCGCCAGTGAACCGTGCACACTCCGGGAGTACATGCGCG 239
Db 76531 AGGCGGGGGAACCTCCGTCGCGCCAGTGAACCGTGCACACTCCGGGAGTACATGCGCG 76590
Qy 240 CGGCGCTGATACCGGCAAGAAATGTCGCAAACTCCCGCTCCGTGAGCGGGCTATTTCAG 299
Db 76591 CGGCGCTGATACCGGCAAGAAATGTCGCAAACTCCCGCTCCGTGAGCGGGCTATTTCAG 76650
Qy 300 GATACCTTTCGTATCATACAGCTACAAACAGAGAGCCAGCTTTTGTCTGACATCCA 359
Db 76651 GATACCTTTCGTATCATACAGCTACAAACAGAGAGCCAGCTTTTGTCTGACATCCA 76710
Qy 360 CAAAGAGGGGAATATTCAAGTCTCGCAGCACTCAACGGCATCGTCAAGTTCGCGCTTGA 419
Db 76711 CAAAGAGGGGAATATTCAAGTCTCGCAGCACTCAACGGCATCGTCAAGTTCGCGCTTGA 76770
Qy 420 ACCCTTAGTATTTTGTCTGTAGTACTATCCAGCAATPAGGTATATCCTGTTGCATC 479
Db 76771 ACCCTTAGTATTTTGTCTGTAGTACTATCCAGCAATPAGGTATATCCTGTTGCATC 76830
Qy 480 AATAAGTTGACTTTGTATCAACATCGGAATTTCCCTTAATCCGGAGCTATTTCGTATG 539
Db 76831 AATAAGTTGACTTTGTATCAACATCGGAATTTCCCTTAATCCGGAGCTATTTCGTATG 76890
Qy 540 ATAAAAAACTCTTCCTGTTCTGATCTCTGCGCTATCGGGAGCTTTTCTACCGCT 599
Db 76891 ATAAAAAACTCTTCCTGTTCTGATCTCTGCGCTATCGGGAGCTTTTCTACCGCT 76950
Qy 600 GTAGCGCTGATAAAAAAGAGACTCAAAATTTCTATATCCAGAAACACTCGATTTAACT 659
Db 76951 GTAGCGCTGATAAAAAAGAGACTCAAAATTTCTATATCCAGAAACACTCGATTTAACT 77010
Qy 660 CCTCTGAGATTACAGAGCCCTGAATCAATCCCTGGGGGCTGATTTTGATATGCCACC 719
Db 77011 CCTCTGAGATTACAGAGCCCTGAATCAATCCCTGGGGGCTGATTTTGATATGCCACC 77070
Qy 720 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAAGATATCAAGATTTTCTGACAACT 779
Db 77071 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAAGATATCAAGATTTTCTGACAACT 77130
Qy 780 TCCAGGATTTGGTCCCTCGGATTTATGGTCATTATGGTCTCTTTCTTTATTCGTATGGCT 839
Db 77131 TCCAGGATTTGGTGGCTCGGATTTATGGTCATTATGGTCTCTTTCTTTATTCGTATGGCT 77190
Qy 840 TGGCAGCGTCCGGAAATACAGGACATATGATGGCCGGGAGGCGCCAGTGGTGGTCAG 899
Db 77191 TGGCAGCGTCCGGAAATACAGGACATATGATGGCCGGGAGGCGCCAGTGGTGGTCAG 77250
Qy 900 CAACGTTTTGAACCGCTGAACAGCTGGCCGAGTAACGTTTAACTCTGATAAAGCCGTCGA 959
Db 77251 CAACGTTTTGAACCGCTGAACAGCTGGCCGAGTAACGTTTAACTCTGATAAAGCCGTCGA 77310
Qy 960 TTGCTGTGGCCAGTCAAGAAAAATPACGGCTCCAGTATTTCTCGGGAGAGCTGATGGTC 1019
Db 77311 TTGCTGTGGCCAGTCAAGAAAAATPACGGCTCCAGTATTTCTCGGGAGAGCTGATGGTC 77370
Qy 1020 CTGACTGTAATGTTGCCCTTGAATCCATGGGATTTAAAAAGCTGGGATTTGCTGGCGGA 1079
Db 77371 CTGACTGTAATGTTGCCCTTGAATCCATGGGATTTAAAAAGCTGGGATTTGCTGGCGGA 77430
Qy 1080 AGAGAAGATGACTGGAGTCGGAACCTGGTATATCTGCGGGGCTGACAAACAGCCCTCTTGCA 1139

Db 77431 AGAGAAGATGACTGGAGTCGGAACCTGGTATATCTGGGGGCTGACAAACAGCCCTTTGCA 77490
Qy 1140 GATAACCGGATAAAAACGGAACTTCAGAAACCTCTTCCGCCACGACGATGGGACTT 1199
Db 77491 GATAACCGGATAAAAACGGAACTTCAGAAACCTCTTCCGCCACGACGATGGGACTT 77550
Qy 1200 ATTTATGTCATCTGAAGCCCGGTGGAACCAAGATCCTCTGGGTTCCCGCAAGAT 1259
Db 77551 ATTTATGTCATCTGAAGCCCGGTGGAACCAAGATCCTCTGGGTTCCCGCAAGAT 77610
Qy 1260 ATCAGGAAGCTTTTTCATGATGCGCATGGAATGAGAGACTGTGGCCTGATCGCG 1319
Db 77611 ATCAGGAAGCTTTTTCATGATGCGCATGGAATGAGAGACTGTGGCCTGATCGCG 77670
Qy 1320 GGAGGGCATACATTTGTTAAAGCACATGTCGACGCTCTCTGAAAAATGTTATTGGCGCA 1379
Db 77671 GGAGGGCATACATTTGTTAAAGCACATGTCGACGCTCTCTGAAAAATGTTATTGGCGCA 77730
Qy 1380 GGGCCTGATGTCACCTGTGGAGGACGAGGACTGGGATGGAATAAATAATGTGTAC 1439
Db 77731 GGGCCTGATGTCACCTGTGGAGGACGAGGACTGGGATGGAATAAATAATGTGTAC 77790
Qy 1440 GGAACCGCAAAATATACCATCAACAGTGGCTCTGGAAGAGCCTGTCGAC 1489
Db 77791 GGAACCGCAAAATATACCATCAACAGTGGCTCTGGAAGAGCCTGTCGAC 77840
RESULT 6
AX191725
LOCUS AX191725 92721 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 7 from Patent WO0149775.
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS
SOURCE
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1
Iversen, P.L.
Antisense antibacterial cell division composition and method
Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source
1. 92721
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 99.2%; Score 1476.4; DB 6; Length 92721;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 1 CTGCAGTCCGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 59
Db 76351 CTGCAGTCCGAGATGAACAGACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCC 76410
Qy 60 ATGATTTTGTTCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC 119
Db 76411 ATGATTTTGTTCACAGACTCAATGACTACCGGACGCACTGAACCTTCGGTGTGTTCTCC 76470
Qy 120 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGTCCGCAAAACCGGGGAAGCATCAGA 179
Db 76471 AGCCAGTTAAAGCCAGCGGTTTCCCTGCTGAAAAATGTCCGCAAAACCGGGGAAGCATCAGA 76530
Qy 180 AGGCGGGGGAACCTCCGTCGCGCCAGTGAACCGTGCACACTCCGGGAGTACATGCGCG 239
Db 76531 AGGCGGGGGAACCTCCGTCGCGCCAGTGAACCGTGCACACTCCGGGAGTACATGCGCG 76590
Qy 240 CGGCGCTGATACCGGCAAGAAATGTCGCAAACTCCCGCTCCGTGAGCGGGCTATTTCAG 299
Db 76591 CGGCGCTGATACCGGCAAGAAATGTCGCAAACTCCCGCTCCGTGAGCGGGCTATTTCAG 76650

Db 76591 CGGCGCTGATACCGCAAGAAATGGTCGCAAACTCCCGCTCCGTGCGAGCGGCTATTTTCAG 76650
Qy 300 GATACCCCTTCCTCATACACAGTCAACACCAAGAGACCGAGCTTTTGTTCGATCCCA 359
Db 76651 GATACCCCTTCCTCATACACAGTCAACACCAAGAGACCGAGCTTTTGTTCGATCCCA 76710
Qy 360 CAAAGAGAGGGAATATTTCAGGTCGCGCAGCACTCAACGGCATCGTCAGTTCGGCGCTTGA 419
Db 76711 CAAAGAGGGAATATTTCAGGTCGCGCAGCACTCAACGGCATCGTCAGTTCGGCGCTTGA 76770
Qy 420 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTGTGCATC 479
Db 76771 ACCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTGTGCATC 76830
Qy 480 AATAAGTTGACTTTCTGATACACATCGCAATTTCCCTTAATCGGAGCTATTTCGTATG 539
Db 76831 AATAAGTTGACTTTCTGATACACATCGCAATTTCCCTTAATCGGAGCTATTTCGTATG 76890
Qy 540 AATAAAAAAATCTTCCTGTTCTGATTTCTCTGGCGCTATCGGGAGCTTTTCTACCGCT 599
Db 76891 AATAAAAAAATCTTCCTGTTCTGATTTCTCTGGCGCTATCGGGAGCTTTTCTACCGCT 76950
Qy 600 GTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTATATCCAGAAACACTGGATTTAACT 659
Db 76951 GTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTATATCCAGAAACACTGGATTTAACT 77010
Qy 660 CCTCTGAGATTACACAGCCCTGATCAATCCCTGGGGGCTGATTTGATTATGCCACC 719
Db 77011 CCTCTGAGATTACACAGCCCTGATCAATCCCTGGGGGCTGATTTGATTATGCCACC 77070
Qy 720 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTTCTGACAACT 779
Db 77071 AGATTTCAACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTTCTGACAACT 77130
Qy 780 TCCAGGATTGGTCCCTGGGATTTATGGTCATTATGGTCTCTTTTATTTCTGATGGCT 839
Db 77131 TCCAGGATTGGTCCCTGGGATTTATGGTCAATATGGTCTCTTTTATTTCTGATGGCT 77190
Qy 840 TGGCAGCGTCCGGAACATACAGACATATGATGGCGGGAGCGCCAGTGGTGTGTCAG 899
Db 77191 TGGCAGCGTCCGGAACATACAGACATATGATGGCGGGAGCGCCAGTGGTGTGTCAG 77250
Qy 900 CAACGTTTGAACCGCTGAACAGCTGGCCGATTAACGTTAATCTGGATAAGCCCGTCGA 959
Db 77251 CAACGTTTGAACCGCTGAACAGCTGGCCGATTAACGTTAATCTGGATAAGCCCGTCGA 77310
Qy 960 TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGATTTTCTGGGAGACCTGATGGTC 1019
Db 77311 TTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGATTTTCTGGGAGACCTGATGGTC 77370
Qy 1020 CTGACTGGTAAATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA 1079
Db 77371 CTGACTGGTAAATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGA 77430
Qy 1080 AGAAGATGACTGGGAGTCCGACCTGTATCTGCGGGCTGACAAAGCCCTCTTGCA 1139
Db 77431 AGAAGATGACTGGGAGTCCGACCTGTATCTGCGGGCTGACAAAGCCCTCTTGCA 77490
Qy 1140 GATAACCGGATAAAAACGGGAACTTTAGAAACCTCTTCCGCGCACCGAGATGGGACTT 1199
Db 77491 GATAACCGGATAAAAACGGGAACTTTAGAAACCTCTTCCGCGCACCGAGATGGGACTT 77550
Qy 1200 ATTTATGTCATCTGTAAGCCCGGTTGAAACACAGATCTCTGTTGGCTTCCGCGAAAGAT 1259
Db 77551 ATTTATGTCATCTGTAAGCCCGGTTGAAACACAGATCTCTGTTGGCTTCCGCGAAAGAT 77610
Qy 1260 ATCAGGGAAGCTTTTTCACCTGATGGCCATGATGAGGAGCTGTGGCCCTGATCGCG 1319
Db 77611 ATCAGGGAAGCTTTTTCACCTGATGGCCATGATGAGGAGCTGTGGCCCTGATCGCG 77670
Qy 1320 GGAGGGCATACATTTGGTAAGACACATGGTGCAGCGTCTCTGAAAAATGATTATGGCGCA 1379
Db 77671 GGAGGGCATACATTTGGTAAGACACATGGTGCAGCGTCTCTGAAAAATGATTATGGCGCA 77730

Qy 1380 GGGCCTGATGTCACCTGTGGAGGAGCAGGACTGGGATGGAAAAATAAATGTGGTACA 1439
Db 77731 GGGCCTGATGTCACCTGTGGAGGAGCAGGACTGGGATGGAAAAATAAATGTGGTACA 77790
Qy 1440 GGAACCGGCAATATACCATCACCAGTGGCTGGAAGGAGCCTGCTGCAC 1489
Db 77791 GGAACCGGCAATATACCATCACCAGTGGCTGGAAGGAGCCTGCTGCAC 77840

RESULT 7
ECEHECCTP
LOCUS ECEHECCTP 2407 bp DNA linear BCT 12-DEC-1996
DEFINITION E.coli 0157:H7 katP gene for EHEC-catalase/peroxidase.
ACCESSION X89017
VERSION X89017.1 GI:1580760
KEYWORDS EHEC-catalase/peroxidase; katP gene; Plasmid.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

1
Brunder W., Schmidt, H. and Karch, H.
KatP, a novel catalase-peroxidase encoded by the large plasmid of
enterohaemorrhagic Escherichia coli O157:H7
Microbiology (Reading, Engl.) 142 (Pt 11), 3305-3315 (1996)
97124214
8969527
REFERENCE 2 (bases 1 to 2407)
Brunder W.
Direct Submission
Submitted (22-JUN-1995) W. Brunder, Universitaet Wuerzburg,
Institut fuer Hygiene und Mikrobiologie, Josef-Schneider-Straesse 2,
D- 97080 Wuerzburg, FRG

FEATURES
source
1. .2407
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="EDL933"
/sub_species="serotype 0157:H7"
/db_xref="taxon:562"
/cell_line="plasmid p0157"
119. .122
131. .2341
/genes="katp"
131. .2341
/genes="katp"
/codon_start=1
/transl_table=11
/product="EHEC-catalase/peroxidase"
/protein_id="CAA61429.1"
/db_xref="GI:1580761"
/db_xref="GOA:P77038"
/db_xref="UniProt/TREMBL:P77038"
/translation="MIKTLPLVLILLALSGSFSTAAVADKKETQNFYYPETLDLTPLR
LHSPENPGADFATRFQQLDMALKKIDKOLLTTTSQDMWADYGHYGFPTIRMAW
HGATVTRTYDGRGSGGQORFELNSPDNVLNLDKARLLNVPVKKGSSISGMDLM
VLTGNVLTSGMFGKTLGAFAGREDWESDLVYMGPNKPLADNRKNGKLGKPLAATQ
MGLIYVNPSPGPGKPDPLASAKDIREAFGRMAMDDEETVALIAGGHTFGKAHGAASPE
KCTIGAGPDGAPVEEQGLGWKNGCTGNGKTYITSGLEGAWSTSPQTFMCKNLKLYK
EWELHSPAGAQWPKKXANIVQDAHDPVLPLMMFTTDIALKVDPEYKKTITREL
NPKATPEAFARAFKLTTHRDMPAARYLGNVEVPASSFIWQDPLPAADYTMIDGKDIK
SLKEQYMDLIGIPASELITKAWASASTFRVTDYRGNGNGARIRLOPEINWENPEKIK
KVLASITSQRFENKKGSDGKVSADLIVLSGNAAIEDAARKAGVELEIPPTPGRTD
ASQEQTDVASFVLEPTADGFRNYYSKRSHTSPVESLIDKASQLDITVPEMTALLGG
LRVMDINTNSISGVFTDTPGVLDNKFVYNLLDMSTWMSKADKEDTYNGFDRKTGALK
WRASSVDLIFSSNPRLRAVAEYASDARNKFHDFVSKWNKWNSDRDFLNNK"
131. .199
/gene="katp"

sig_peptide
131. .199
/gene="katp"

ORIGIN

Query Match 72.6%; Score 1081.4; DB 1; Length 2407;
Best Local Similarity 99.9%; Pred. No. 6.5e-311;

		Matches 1082; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	407	GTTCGGCTTGGAACCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTAT	466		
Db	1	GTTCGGCTTGGAACCCCTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTAT	60		
Qy	467	ATCCTGTGCATCAATTAAGTTGACATTTGTATACAACTGCGAATTTCCCTTAATCCGG	526		
Db	61	ATCCTGTGCATCAATTAAGTTGACATTTGTATACAACTGCGAATTTCCCTTAATCCGG	120		
Qy	527	AGCTATTTCGTATCAAAAAAATCTCTCTGTCTGTAGTATCTTCTGGCGTATCGGGAG	586		
Db	121	AGCTATTTCGTATCAAAAAAATCTCTCTGTCTGTAGTATCTTCTGGCGTATCGGGAG	180		
Qy	587	CTTTTCTACCGCTGTAGCCCTGATAAAAAGAGACTCAAAATTTCTATCTACAGAAC	646		
Db	181	CTTTTCTACCGCTGTAGCCCTGATAAAAAGAGACTCAAAATTTCTATCTACAGAAC	240		
Qy	647	ACTGGATTAACTCCTCTGAGATTACACAGCCCTGATCAAAATCCCTGGGGGCTGATTT	706		
Db	241	ACTGGATTAACTCCTCTGAGATTACACAGCCCTGATCAAAATCCCTGGGGGCTGATTT	300		
Qy	707	TGATTATGCCACCAAGATTTCAACAGCTGGATATGGAGCTCTCAAAAAAGATATCAAGA	766		
Db	301	TGATTATGCCACCAAGATTTCAACAGCTGGATATGGAGCTCTCAAAAAAGATATCAAGA	360		
Qy	767	TTTGTCTGACAACTTCCAGGATTTGGTCCCTGCGGATATGGTCAATATGGTCTTTCTTT	826		
Db	361	TTTGTCTGACAACTTCCAGGATTTGGTCCCTGCGGATATGGTCAATATGGTCTTTCTTT	420		
Qy	827	TATTCTGATGGCTGGCAGCGTCCGGAACATACAGGACATATGATGCGCGGGAGCGC	886		
Db	421	TATTCTGATGGCTGGCAGCGTCCGGAACATACAGGACATATGATGCGCGGGAGCGC	480		
Qy	887	CAGTGGTGGTCAAGCAAGCTTTTGAACCGCTGAAACAGCTGGCCGGATTAACCTCTGGA	946		
Db	481	CAGTGGTGGTCAAGCAAGCTTTTGAACCGCTGAAACAGCTGGCCGGATTAACCTCTGGA	540		
Qy	947	TAAAGCCCGTCGATTTGCTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGG	1006		
Db	541	TAAAGCCCGTCGATTTGCTGGCCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGG	600		
Qy	1007	AGACCTGATGGTCTGATGTTATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGG	1066		
Db	601	AGACCTGATGGTCTGATGTTATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGG	660		
Qy	1067	ATTTGCTGGCGAAGAGATGACTGGAGTCCGACCTGTGTATCTGGGGCTGACAA	1126		
Db	661	ATTTGCTGGCGAAGAGATGACTGGAGTCCGACCTGTGTATCTGGGGCTGACAA	720		
Qy	1127	CAAGCCTCTTGCAGATAACCGGGATAAAACGGGAACTTTCAGAACTCTTGGCGGCAC	1186		
Db	721	CAAGCCTCTTGCAGATAACCGGGATAAAACGGGAACTTTCAGAACTCTTGGCGGCAC	780		
Qy	1187	GCAGATGGGACTTATTTATGTCATCTGAGGCCCGGTGGAAACAGATCCCTCTGGC	1246		
Db	781	GCAGATGGGACTTATTTATGTCATCTGAGGCCCGGTGGAAACAGATCCCTCTGGC	840		
Qy	1247	TTCCGCGAAGAGATATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGCTGT	1306		
Db	841	TTCCGCGAAGAGATATCAGGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGCTGT	900		
Qy	1307	GGCCCTGATCCGGGAGGCAATACATTTGGTAAAGCAATGGTCAGCGTCTCTGAAAA	1366		
Db	901	GGCCCTGATCCGGGAGGCAATACATTTGGTAAAGCAATGGTCAGCGTCTCTGAAAA	960		
Qy	1367	ATGTATTGGCCGAGGCTGTATGTCACCTGTGGAGGAGCAGGACTGGGATGAAAAA	1426		
Db	961	ATGTATTGGCCGAGGCTGTATGTCACCTGTGGAGGAGCAGGACTGGGATGAAAAA	1020		
Qy	1427	TAAATGTGGTACAGGAACCGCAATATACCATCACAGTGGCCTGGAGGAGCCTGGTC	1486		
Db	1021	TAAATGTGGTACAGGAACCGCAATATACCATCACAGTGGCCTGGAGGAGCCTGGTC	1080		

Qy 1487 GAC 1489
Db 1081 GAC 1083

RESULT 8
BX936398_09/c
WPCOMMENT

Sequence split into 48 fragments LOCUS BX936398 Accession BX936398

Fragment Name	Begin	End
BX936398_00	1	110000
BX936398_01	100001	210000
BX936398_02	200001	310000
BX936398_03	300001	410000
BX936398_04	400001	510000
BX936398_05	500001	610000
BX936398_06	600001	710000
BX936398_07	700001	810000
BX936398_08	800001	910000
BX936398_09	900001	1010000
BX936398_10	1000001	1110000
BX936398_11	1100001	1210000
BX936398_12	1200001	1310000
BX936398_13	1300001	1410000
BX936398_14	1400001	1510000
BX936398_15	1500001	1610000
BX936398_16	1600001	1710000
BX936398_17	1700001	1810000
BX936398_18	1800001	1910000
BX936398_19	1900001	2010000
BX936398_20	2000001	2110000
BX936398_21	2100001	2210000
BX936398_22	2200001	2310000
BX936398_23	2300001	2410000
BX936398_24	2400001	2510000
BX936398_25	2500001	2610000
BX936398_26	2600001	2710000
BX936398_27	2700001	2810000
BX936398_28	2800001	2910000
BX936398_29	2900001	3010000
BX936398_30	3000001	3110000
BX936398_31	3100001	3210000
BX936398_32	3200001	3310000
BX936398_33	3300001	3410000
BX936398_34	3400001	3510000
BX936398_35	3500001	3610000
BX936398_36	3600001	3710000
BX936398_37	3700001	3810000
BX936398_38	3800001	3910000
BX936398_39	3900001	4010000
BX936398_40	4000001	4110000
BX936398_41	4100001	4210000
BX936398_42	4200001	4310000
BX936398_43	4300001	4410000
BX936398_44	4400001	4510000
BX936398_45	4500001	4610000
BX936398_46	4600001	4710000
BX936398_47	4700001	4744671

Continuation (10 of 48) of BX936398 from base 900001 (BX936398 Versinia pseudotuberculosis

Query Match 36.3%; Score 539.8; DB 1; Length 110000;
Best Local Similarity 69.9%; Pred. No. 8.9e-149;
Matches 742; Conservative 0; Mismatches 317; Indels 2; Gaps 1;

Qy	429	TATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTTCATCAATAAGTT	488
Db	75176	TATCAATAAGTAGAATACTTATCATATTAATAGGACTTAACGAATTTGTAATTTCCAA	75117
Qy	489	GACTTTGTATACACATGCGAATTTCCCTTAATCCGAGCTATTTCGTATCATAAAAA	548
Db	75116	AACCATGTTAAAAAATAGTTGCCCTTTTAGTTAAAGGGGACTTAT--ATATGTTAAAAA	75059


```
/translation="MNINLMVNIYSMRRKIMKNTLLKSLCTAALISMAGVAGAASNGL
IAILTPSHNPFFKAEAGAKATELGYTTLIVASHDDVDNKNQOLIEITAIARKAKAI
ILDNAGSDATIGLKKAKAGIPAFILIDREINETGIAVSIQVSNVYQAGLGAKEPVT
LMGKGYTELIGRESIDNAHVRSGYIHVDIHEHSDMKVAAQQTANWSTQTEAFNRHES
ILQANPDITGVISGNDTMAIGABAAALKAARNDVVVVGFDGSDYVDRDSIINNGTKAT
VLQPGWAQAAVAVQADQYLTKTGKTEGSKQLMDCVLDENNKNLNVFALKD"
3893..4510
/gene="y0872"
CDS
3893..4510
/function="unknown"
/notes="residues 1 to 205 of 205 are 100.00 pct identical
to residues 1 to 205 of 205 from GenPept :
>gb|AAD37311.1|AF135170_2 (AF135170) unknown [Yersinia
pestis]"
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AAM84457.1"
/db_xref="GI:21957607"
/translation="MRPRAISKAGALLMSAVLVACTVVLDLNGKPIPLVDPNAVVS
DINQDPDKVAITIAKWNPPANSNAISHQEVKQSQSPAVGKNQSRFVRTGKVA
VTEGREGILRIALIDGDEQVLQGPVIVKGNARIDASTFIRFDDFKNQVQYQALAKLS
LRALQVAKPDASWIGQVEVLAAVTLTPTLSNGVPLSLNKESH"
4510..6033
/gene="y0873"
CDS
4510..6033
/function="putative transport"
/notes="sugar (D-ribose?) transport; residues 1 to 344 of
507 are 100.00 pct identical to residues 1 to 344 of 344
from GenPept : >gb|AAD37310.1|AF135170_1 (AF135170) ribose
ABC transporter [Yersinia pestis]"
/codon_start=1
/transl_table=11
/product="putative ATP-binding protein of ABC transport
system"
/protein_id="AAM84458.1"
/db_xref="GI:21957607"
/translation="MNRHFDIVNRADISMRFPFGTLALDSVNTVVRGKVVIGENG
AKGKTLMLTAGVQQTQGIYLNQKVSIAINTREAAALGIVMHOELNSLENLNAE
NIFLGRKIEIQLKPNQAAQRETEQLMRLDQVISPQKSNLKVQOQQLIEIAKAL
RQADILIDEPTEALSKEVDVLFVIRELTRQGVSIIVISHRLEELMAGDYITIL
RGRFOAEAAVDIDVPMIVREMGSDPVSSFLRPGRTFGBPMVEVDVTLNESGNT
VNVVSLHVRAGEIIVIGLMNGRTELECELLGTYQNYLGTVRLANGTVMNAKTAE
RARFGMSLVPDRKKAQIPASSVACNLISLWRKLKRFALWQBSVEVQVVAATIGE
LSIKVSSPDVIEQALSGNQKQKVIGRSLLTRPNILLDEPTRGIDVGAKADVEMVY
KLSEQGIGLFTSDLKEIMAVSDRILVMSNGKLTANLSRQSASESALVTASAQGPE"
6030..7175
/gene="rbsC"
/notes="synonym: y0874"
6030..7175
/gene="rbsC"
/function="transport; transport of small molecules;
carbohydrates, organic acids, alcohols"
/notes="residues 1 to 356 of 381 are 59.83 pct identical to
residues 1 to 348 of 354 from GenPept : >emb|CAC48738.1|
[Au03643] probable sugar ABC transporter permease protein
[Sinorhizobium meliloti]"
/codon_start=1
/transl_table=11
/product="sugar (D-ribose?) transport system permease
protein"
/protein_id="AAM84459.1"
/db_xref="GI:21957608"
/translation="MNKATGTALASHLPFGASWSRENMLLLLMKRTFIALFLILGF
FSIMVPGFLATSLIIMVHIAINAFALGIVFIITAGIDLSIGATLGLCGMIAGWM
ITQIGVLPWFEGATIPSVVWVPLVIGALIGAVNGWIIITRYNAPPTCTLGTWVYV
RGASMLISGEPFPGLOGNPGTGLGSLGLGLPIAWIMFILAIYAVARR
LPFGRHVYAIAGNERRAELSGVKVRQVWVVTLSGFCAAIAGIIVSAQLVASHPANG
SGFEMNAIAVNGTSGLAGRGITLIGTIGAPVILGADGLVMGMSISFQWQVINGI
VIIVAVIIDQMNRMQHKAAIVSQKVTAPSDVTSQAGIISQTVKSE"
7426..9193
/gene
```

```
/gene="y0875"
/notes="disrupted by frameshift"
/pseudo
9196..10683
/gene="glpK"
/notes="synonym: y0876"
9196..10683
/gene="glpK"
/function="enzyme; central intermediary metabolism: Pool,
multipurpose conversions"
/notes="residues 6 to 490 of 495 are 40.36 pct identical to
residues 7 to 494 of 502 from E. coli K12 : B3926;
residues 6 to 490 of 495 are 44.35 pct identical to
residues 4 to 479 of 489 from GenPept : >gb|AAL42899.1|
(AE009144) glycerol kinase 2 [Agrobacterium tumefaciens
str. C58 (U. Washington))"
/codon_start=1
/transl_table=11
/product="glycerol kinase"
/protein_id="AAM84460.1"
/db_xref="GI:21957609"
/translation="NSTPIILAIIDSGTTNAKAIANDRAGRVLAKASVPLQLEHPQGM
AEQDPLAIWLAVSQAVEGCLNQAQVAGIAISNQRSVLIERWTGTPLPFVVSQWQ
CRRSAFCLALRQLPAAAVABERTGLIDPLPAAKHGMLAQIIPHGAARAHEGLCV
GTLDCWLTFQFSEGSFSTDFNAARTQLDNHSGQMDPLLALFGLPCLPFAVLPS
ASIHGHTARTGILGLEQGVPIVAHIGDSHAALYGQGDQAGIKATYGTGSSLMVTLK
QASAPSHGLSTTIAMHDELRYALRGNTHTGSGVAVWSRMLGISLRLTDMABSQS
GNQGVYFPALSGLGPAYDSQARGLFCGLTDATTPAVLARAALGESVAQIADVFAM
ENAAHRLERLRVGGATSNRMWQFQADLLORTLIRNHTAEVSALGAAYLGGKTLGW
WNAQDUALALPREVEYIEPRIHSAEMQDNYSLWQTAIARARFQPK"
10704..12128
/gene="y0877"
CDS
10704..12128
/function="unknown"
/notes="residues 7 to 471 of 474 are 51.28 pct identical to
residues 3 to 467 of 471 from GenPept :
>gb|AAD36032.1|AE001758_10 (AE001758) hypothetical protein
[Thermocoga maritima]"
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AAM84461.1"
/db_xref="GI:21957610"
/translation="NSALPRKLIKFAVIIGNRGPFPFSLVSDARRDAQALFDRLGINII
MLNEQTPLGGVNHWDKAKVCAELRQHSSEIHGVVVLNPFNGDEKSVSDAIRLSLK
VPLVQABEDALDKMLATRRDSPCKISLKNLRQYGIPTLITKOHVCSLMSSEISFR
DVEQVQVLCRVVHAMKRVVGAIGARPNTFNTVRYSEKLENLGITVETLIDUSEVFFR
Query Match 36.0%; Score 536.6; DB 1; Length 12268;
Best Local Similarity 69.7%; Pred. No. 4.7e-148;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;
Qy 429 TATTTTTCCTGTGTAGTATCTATCCAGCAATAGGTATATCTGTGTGCATCAATAAGTT 488
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2429 TATCAATAATAGATAGAACTCATATTAATAGGACTAACGAATTTGTTATTTTCCAATTT 2370
Qy 489 GACTTTTGTATACACATCGGAATTTCCCTTAATCCGGAGCTATTTCGTATGATAAAAAA 548
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2369 AATCATGTTTAAAAAATAGTTGCTTTTGTAGTAAAGGGGACTTAT--ATATGTTAAAAAAA 2312
Qy 549 ACTCTTCTCTGTTCTGATCTCTTGGCGCTATCGGGGAGCTTTTCTACCGCTGTAGCGGCT 608
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2311 ATCTTACCGGTACTATACTCTCGGCATTTGATATAATACACTACGGCTTGGGCAGCA 2252
Qy 609 GATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGATTTAACTCTCTTGAGA 668
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2251 GAAGCACCCAAGACTGATAGTTTTTACTTACCCAAAAGCCTCGATCTCTCTCCACTCCGT 2192
Qy 669 TTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGGATATATGCCACAGATTTCAA 728
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2191 TTACATAATATAGAGTCCAATCCTTACGAAAAGGATTTTAAATTATGCTCAACAGATTAA 2132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

MGLIYVNPENGPNGKDPVAAAKDIREAPARMAMNDBETVALIAGGHTFGKAHGAASPE
KCLGAAPGEAGLEQOGLGWANKSGNGKDTITSGLEGAWTDTPTHTMYOYLNLKYKH
EWLWTSKPAGANQWOKPKNAANVVPDADTDPTKPHLMMFTDIALKVDPEYKKITRFL
ENPEEKSPAGAWFAXLITKMDMGPAARYLGDVEPKETFIWDDPLPAANYKMDISADIS
ELKDKLTKGLSKTKLITKAWASAFRFGDPRGNGGARIRLAPQKQVNDVPPAEHL
SVLAALMEVONNFKNORSCKVYSLDLIVLGNAAIEDAACKAGYSISIPFTPGRTD
ASQEBTDVSSFAVLEPTADGFRNYDAKRNLTSPILASLIDRANKLELTPVEMTLVIGG
LRVLVDVSSGSKAGVLNTPGQNNPFVNLLDMSTKWTSPKAEVGFQGVDRKTKGL
KWTASSVDLVFGSNPELRAVEVYASDDAKEKFHDFTKVKWEKMLDRFDIKN"
5511..5894
/gene="cybC"
CDS
5511..5894
/gene="cybC"
/note="similar to Escherichia coli cytochrome c-562"
/transl_table=1
/codon_start=1
/product="cytochrome c"
/protein_id="AAD37314.1"
/db_xref="GI:5002123"
/translation="MRKPIIAGVFSLLITSCTFASLQDDMNLILANLIGVSSSTD
KVLITSSLEKRNAAALDQKALPPKLEGAEDSPKIDYRNGFDLLIEQIDKTKQWABE
GNIQEVKKSVEGINRTNTHSYR"
6042..6572
/gene="cybB"
CDS
6042..6572
/gene="cybB"
/note="similar to Escherichia coli cytochrome b-561"
/transl_table=1
/codon_start=1
/product="cytochrome b"
/protein_id="AAD37315.1"
/db_xref="GI:5002124"
/translation="MHKYSCLQIGIHWLVLLVLIWSSIELRGFAPRSGYQWMMI
HFSCGIALIWMTRILIQRYPTPIVPKPSMIVGLAHVGHVYLLPFIPIGI
AIIYCRGSSWIAFGLIMHAEQANFDLADTLKAYHLLANNVFPVIGLHALALHHY
VLKDNLTLRWMPKRG"
complement(6829)..7539
/note="ORF7; hypothetical protein"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="unknown"
/protein_id="AAD37316.1"
/db_xref="GI:5002125"
/translation="MIKIDMTYQDIYVKDYQDSVYQKILGTRVLGALFYCEPDSA
ECRDIYVOLRNGTWAAEPYGGADELPIATLLAAQSDQYQSIHAQSAAPLAETLE
EAWQLRFGYALPAPWGSYLDKDMVLFGDSTLKRNMWQQRHVETLKQOEPEDH
FGLIMMAWAAAEHQVVDLPVLLADHLLPWSRYLALLQSDAGHPFYQGLAQLTTL
AHWQHELQVTPAGVELYR"
complement(7685)..8545
/gene="dmsC"
CDS
complement(7685)..8545
/gene="dmsC"
/note="similar to Escherichia coli DMSO reductase"
/transl_table=1
/codon_start=1
/product="dimethyl sulfoxide reductase subunit C"
/protein_id="AAD37317.1"
/db_xref="GI:5002126"
/translation="MGSQHEWPLMWFTVLQCVVGGVIFVLGLALIFGGLSRGQOQRV
HRSNFWLWLMALAFIATSLHLSGPMRAFNSLRIGESALSNEIAGSLPFAVAGFYW
LLAVLDKMPAALGKIWLIVANVLGVIFVYAMCRYYSINTVPTWDSIYTPILGFVLTLL
GGPILGGLYLQADVNGRAMIQLEPMISVLALIIISVASVYMQAASIPLIYSSVQASAL
IPDYGLSMWRLVLLVLGLGCWICPLIRGRPMPLVMVLGTLVLVAGELIRGVFYGL
HMTVGMVAVGG"
complement(8547)..9164
/gene="dmsB"
CDS
complement(8547)..9164
/gene="dmsB"
/note="similar to Escherichia coli DMSO reductase"
/transl_table=1
/product="dimethyl sulfoxide reductase subunit B"

```

```

/protein_id="AAD37318.1"
/db_xref="GI:5002127"
/translation="MTQYGFYIDSSRCKTCKTELACKDFKNLSPDVSPRRYIYEYAG
GWQOQNGTWHQNVFAYVILSIACNHCSDPACTKVCPSGAMHKRDDGGVNVNBDICIG
RYCHMACPAQFOYDAEKHMTKCDQCYERVA/GKKPICVDSCELRALDWPIDELRE
KYGDIAEIAPLPAAHFTLNPVILKPNANRSPVRDITTHGLANPKEV"
complement(9176)..11647
/gene="dmsA"
CDS
complement(9176)..11647
/gene="dmsA"
/note="similar to Escherichia coli DMSO reductase"
/codon_start=1
/transl_table=1
/product="dimethyl sulfoxide reductase subunit A"
/protein_id="AAD37319.1"
/db_xref="GI:5002128"
/translation="MQLSGNLMKETNESRLITGVTRKLVQVTLVGLAMATCAPSL
PFSRTARAVQSALNPTSGANDKVIWISACTVNCGRSCLPRLMHVADGEIKYVETONTGDD
DFEGLHQVRACLRGRSMRRVYNPDLKYPMKRIGARGEGKFKRISWEEAFETIAGSM
QNIITKEYGNEAIYLYGTGTIGTMTSRSPPGATLLARLMCCGGYLYLEQAREKSNKM
AAGLNTYCGWADGNSPDSIENSKLVLFGNPAETRMSCGGVYVYLEQAREKSNKM
IYDPRYDTAAGREDEWIPLRGTDAAALASALAYVMTENLVDQPPDOYCVGYDEK
TLDPGAPENSHYKAYILQGGADAKTPQWAAALITGIPAAKIITLREISVRPAYIS
QGWGPQRSHNGENTSAIALMLAILTNGVINGNGSAREGSYLPFVFMPTFEFPVKI
SISMFLWTDAIERGPEMTATRDGVQKDLDPVKFTWYASNLNQHSEINRTHDI
LODDKCEMIWIDNHTWSAAYADIILPDCTASEQMDFLDASASGNMAYVIFADQVI
KPRFECKNIDITSLGARHMGVEQTEGTRQEBEMRHLVQOQAAIPELSPSFDPR
QGMFKRDPAGHVAVHDFRDPVAMPLETPSPKIEIYSAALAEIATWELQKDDVID
PLPYAAGPESHVDPLIEQYPLQITGFHYKARTHSTYGNVDVLKASCLQEWLNPMDA
ERRGISNGDMVRIFGNGEIVINAKVTFRMPMPGVVALGEGVAVSPDAKRIDRAGSINV
LTTQRPSPLAGNPSHTNLVQVAKA"
complement(12089)..12859
/gene="araD"
CDS
complement(12089)..12859
/gene="araD"
/note="similar to Escherichia coli ribulose epimerase"
Query Match 36.0%; Score 536.6; DB 1; Length 13632;
Best Local Similarity 69.7%; Pred. No. 4.9e-148;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;
Qy 429 TATTTTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTCTTGGATCAATAAGTT 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3115 TATCAATAATAGTAGAATCTATCATATTAATAGACTAACGAATGTATTTTCCAAATTT 3174
Qy 489 GACTTTTGTATACACATCGGAATTCCTTAATCCGAGCATTCGTATCATATAAAAAA 548
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3175 AATCATGTTAAAAAATAGTTGCCCTTTTAGTTAAAGGGGACTTAT--ATATGTTAAAAA 3232
Qy 549 ACTCTCTCTGTTCTGATTTCTTGGCGCTATCGGGGAGCTTTTCTACCGCTGTAGCCGCT 608
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3233 ATCTTACCGGTACTAATACTCTGCCATTTGATACATAATACACCTACGGCTTGGGCAGCA 3292
Qy 609 GATAAAAAAGAGACTCAAAATTTTCTACTATCCAGAAACACTGGATTTAACTCTCTGAGA 668
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3293 GAAGCACCAAGACTGATAGTTTTTTACTACCCAAAAGCCCTCGATCTCTCTCCACTCCGT 3352
Qy 669 TTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGATTTATGCCACAGATTTCAA 728
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3353 TTACATAATATAGAGTCCAATCTCTACGGAAGGATTTTAATATGCTCAAACAGTTTAA 3412
Qy 729 CAGCTGATATGGAGGCTCTGAAAAAAGATATCAAGATTTTGTCTGACAACTTCCAGGAT 788
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3413 ACATGGATCTTGAAGCAGTAAAGAGATATAAACAAGTCTTACCATCAACAGAC 3472
Qy 789 TGGTGCCCTGCGGATTATGGTCATTATATGGTCCTTTCTTTTATTCGTATGGCTTGGCAGGT 848
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3473 TGGTGGCTGCTGATTATGGTAAATATATGTCCTTCTTTTATTCGTATGGCTTGGCATGGT 3532
Qy 849 GCCGGAACATACAGACATATGATGGCGGGGCGCGCCACTGGTGGTGCAGCAAGCTTTT 908
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3533 GCGGGAACCTACCCGCATATATGATGGCGGTGGTGGTGGCGGCGAGCAAGATTTT 3592

```

Qy	909	GAACCGCTGAACAGCTGGCCGAGTAACCTTAATCTGGATAAAGCCGCTGATTCGTGTGG	968
Db	3593	GAGCAGCTCATAGCTGGCCAGATACCCACCTTGATAAAGCCGCTCGGCTTCGTGG	3652
Qy	969	CCAGTCAAGAAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGGTCTCTGACTGGT	1028
Db	3653	CCTATAAAAAAGAAATATGGCGCTAAAAATATCGTGGGCGACCTGATGGTCTCTTACAGGC	3712
Qy	1029	AATGTTGCCCTTGAATCATCGGATTTAAACCGTGGGATTTGCTGGCGGAGAGAGAGAT	1088
Db	3713	AATGTCGGCTCGAATCTATCGGGTTTAAAAACGCTGGGGTTTGCAGGAGGCGCTGAGGAT	3772
Qy	1089	GACTGGGAGTCCGACCTGGTATCTAGTGGGGCTGACAAACAGCCCTTTGCAGATAACCGG	1148
Db	3773	GATTGGCATCCGATCTGCTACTTGGGGCGAGGCAACAAGATGCTCTCTGATAACAGG	3832
Qy	1149	GATAAAAACGGAAACCTTCAGAAACCTCTTCCCGCACGAGATGGGACTTATTATGTC	1208
Db	3833	GATAAAATGCAAGCTTACCCAAACCGCTGGCGGCACACAAATGGGGTTGATTACGTA	3892
Qy	1209	AATCCTGAGCCCGGTGGAAACCCAGATCTCTGGCTTCGGCGAAAGATATCAGGAA	1268
Db	3893	AACCCCTGAAGACCAAAATGGTAACCCAGACCCGGTTGCTGGCGCAAAAGATATTCGTGA	3952
Qy	1269	GCTTTTTCACGTATGGCATGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGGCAT	1328
Db	3953	CGTTTGGCAGATGGCANTGAATGATGAGGAGACCGTGGGATGATTCGGGAGGGGAT	4012
Qy	1329	ACATTGGTAAAGCACATGGTCGACGCTCTCTGAAAATGATTTATGGCGAGGGCCTGAT	1388
Db	4013	ACCTTCGTAAAGCGCATGGCGACGCTCTCTGAAAATGCTGCTGGTGACACCGGGT	4072
Qy	1389	GGTCACCTGTGGAGGACGAGGACTGGGATGGAAAAATAATGTGGTACAGGAAACGGC	1448
Db	4073	GAACGGGGCTAGAACCAACAGGTTTAGTAGTGGCGGATAAATGTGTTACAGTAATGGT	4132
Qy	1449	AAATATACCATCACAGTGGCTGGAGGACCTGCTGCAC	1489
Db	4133	AAGACACGATACCACTGGTGGTGGAGGGGATGGACAC	4173
RESULT 11			
AJ414156			
LOCUS			
Yersinia pestis strain CO92 complete genome; segment 16/20.			
AJ414156 AL590842			
AJ414156.1 GI:15981150			
KEYWORDS			
Yersinia pestis CO92			
ORGANISM			
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
Enterobacteriaceae; Yersinia.			
REFERENCE			
AUTHORS			
1 (bases 1 to 220050)			
Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T.G.,			
Prentice, M.B., Sebahia, M., James, K.D., Churcher, C., Mungall, K.L.,			
Baker, S., Basham, D., Bentley, S.D., Brooks, K., Cerdono-Tarraga, A.M.,			
Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G.,			
Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Leather, S.,			
Karllyshev, A.V., Moule, S., Oyston, P.C.F., Quail, M., Rutherford, K.,			
Slimmons, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G.			
Genome sequence of Yersinia pestis, the causative agent of plague			
Nature 413 (6855), 523-527 (2001)			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
2 (bases 1 to 220050)			
Parkhill, J.			
Direct Submission			
TITLE			
JOURNAL			
Submitted (04-OCT-2001) Submitted on behalf of the Yersinia			
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,			
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk			
Notes:			
Details of Y. pestis sequencing at the Sanger Centre are available			
on the World Wide Web.			

FEATURES		(URL, http://www.sanger.ac.uk/Projects/Y_pestis/).
source		Location/Qualifiers
		1..220050
		/organism="Yersinia pestis CO92"
		/mol_type="genomic DNA"
		/strain="CO92"
		/db_xref="taxon:214092"
		/note="biovar: Orientalis"
		/complement(191..715)
		/gene="arol"
		/note="synonym: YPO3215"
		/complement(191..715)
		/gene="arol"
		/EC_number="2.7.1.71"
		/note="Similar to Escherichia coli shikimate kinase II
		AroL SW:AROL_ECOLI (P08329) (174 aa) fasta scores: E():
		1.5e-31, 55.4% id in 168 aa, and to Erwinia chrysanthemi
		shikimate kinase AroL or AroM SW:AROL_ERWCH (P10880) (173
		aa) fasta scores: E(): 0, 63.7% id in 168 aa"
		/codon_start=1
		/transl_table=11
		/product="shikimate kinase II"
		/protein_id="CAC92450.1"
		/db_xref="GI:15981151"
		/db_xref="GOA:Q8ZC15"
		/db_xref="UniProt/Swiss-Prot:Q8ZC15"
		/translation="MTQTFVVGARGAGKTTIGKALAAQALGYRFVDTDLFMQOTSQMT
		SAEVSEGDWDFRLRESMALQAVTAPKTVATGGGAVLSSENAFMRDHGRVYIYLA
		AAVLAARLAEADPEEAQRPSTLTKPIVEILDVLAAREALYQDVAHVLDGTQTPTSLV
		VEQILQMLTGEMVK"
		/complement(209..709)
		/gene="arol"
		/note="Pfam match to entry PF01202 SKI, Shikimate kinase,
		score 245.00, E-value 1.1e-69"
		/complement(470..544)
		/gene="arol"
		/note="PS01128 Shikimate kinase signature."
		/complement(668..691)
		/gene="arol"
		/note="PS00017 ATP/GTP-binding site motif A (P-loop)." "
		/complement(962..1519)
		/gene="YPO3216"
		/complement(962..1519)
		/gene="YPO3216"
		/note="Similar to Rhizobium meliloti phospholipid
		N-methyltransferase PmtA TR:AAG10237 (EMBL:AF201699) (200
		aa) fasta scores: E(): 1.1e-07, 25.9% id in 185 aa, and to
		Bradyrhizobium japonicum PmtA protein TR:Q9LCT2
		(EMBL:Y09633) (199 aa) fasta scores: E(): 0.0057, 24.2% id
		in 178 aa, and to Klebsiella pneumoniae hypothetical 22.7
		kDa protein TR:O87755 (EMBL:AJ011907) (196 aa) fasta
		scores: E(): 0, 52.9% id in 191 aa"
		/codon_start=1
		/transl_table=11
		/product="putative methyltransferase"
		/protein_id="CAC92451.1"
		/db_xref="GI:15981152"
		/db_xref="GOA:Q8ZC14"
		/db_xref="UniProt/TTrEMBL:Q8ZC14"
		/translation="MLNIVRNPKSOFYLQRFATSPRTVGSLSVSPSPWLCOAMLNQI
		DWTQNLNIAELGADGVLTGRILSHMSANSLSQAYEIQPHFVHALHQLKDSRLQIADR
		SABQLODEYDVVFCCLPLLSIPTKISIRILQOQTLRARNRNGVLVLFQYSHLSBSLLS
		RYFNKKIRVVRNPPALVYICHFN"
		/complement(1787..2002)
		/gene="YPO3217"
		/complement(1787..2002)
		/gene="YPO3217"
		/note="no significant database hits"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="CAC92452.1"
		/db_xref="GI:15981153"


```
/db xref="Uniprot/T:EMBL:Q8ZC13"
/translation="MGITDVGIVVANVIMAEITNDRGRSISCHFFIQSSPNAHLNVD
IKNIASPHVLSIVACSIKKAARLRUS"
complement (2249. .2770)
/gene="yPO3218"
CDS
complement (2249. .2770)
/gene="yPO3218"
/notes="Similar to Escherichia coli hypothetical protein
YaFe SW:YAFB_ECOLI (P30866) (207 aa) fasta scores: E(): 0,
65.1% id in 169 aa, and to Pseudomonas aeruginosa YaFe or
PA3119 TR:O87011 (EMBL:U93274) (187 aa) fasta scores: E():
0, 55.2% id in 172 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92453.1"
/db xref="GI:15981154"
/db_xref="GOA:Q8ZC12"
/db_xref="UniProt/T:EMBL:Q8ZC12"
/translation="MQAAADKLTNIEIQGLAESLPADQSPDIVISRYSAHHWHDV
GKALREKRVLPFGKVFMDVSPGHPVLDIYQIVELRDTSHVRNYSFGWLSLF
TEGLVINEVTSRDLYLEFSWARMETPAHFATAIREFOKLASDGVINHAIQADGS
FTSDIMWIVARD"
complement (3016. .3285)
/gene="yPO3219"
complement (3016. .3285)
/gene="yPO3219"
/notes="no significant database matches"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAC92454.1"
/db xref="GI:15981155"
/db_xref="UniProt/T:EMBL:Q8ZC11"
/translation="MEGSVMGISEESIRRLTNEKNVGHSAKWVAIISAYFIIMLF
YKHELGLVLTAGGFLVFSFTTMMKQKVKSYKNLQOIEBKT"
complement (3088. .3138)
/gene="yPO3219"
/notes="one of 2 probable transmembrane helices predicted
for yPO3219 by TMHMM2.0"
complement (3151. .3216)
/gene="yPO3219"
/notes="one of 2 probable transmembrane helices predicted
for yPO3219 by TMHMM2.0"
complement (3468. .4226)
/gene="yPO3220"
complement (3468. .4226)
/gene="yPO3220"
/notes="Similar to Escherichia coli hypothetical protein
YcGR SW:YcGR_ECOLI (P76010) (244 aa) fasta scores: E(): 0,
40.7% id in 241 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92455.1"
/db_xref="GI:15981156"
/db_xref="UniProt/T:EMBL:Q8ZC10"
/translation="MGPVSETSKQFVKKNKLAICAILRDLQKNDTAVMVTARGQFI
SRILDIQETNQIFDFGSVENENVALGAECQLTIIVEPTGAKIEFTCNKLHVYELS
LPFSSAIPLOLYTORREYFVSIPOWYVCSGKPPDQYKYVTLADIISLGMGLY
AMKSEPLQCSVLRODAVDLQGLGFLKDLQFLRALDKQVNNKGEMLTVQRLSPK
FLRUSPQERGLQRAIFELERQQTAKARKFQGL"
complement (4554. .5813)
/gene="proA"
/notes="synonym: YPO3221"
complement (4554. .5813)
/gene="proA"
/EC_number="1.2.1.41"
/notes="Similar to Serratia marcescens gamma-glutamyl
phosphate reductase ProA SW:PROA_SERMA (P17857) (417 aa)
fasta scores: E(): 0, 85.1% id in 417 aa, and to
Escherichia coli gamma-glutamyl phosphate reductase ProA
SW:PROA_ECOLI (P07004) (417 aa) fasta scores: E(): 0,
```

```
72.4% id in 417 aa"
/codon_start=1
/transl_table=11
/product="gamma-glutamyl phosphate reductase"
/protein_id="CAC92456.1"
/db xref="GI:15981157"
/db_xref="GOA:Q8ZC09"
/db_xref="UniProt/Swiss-Prot:Q8ZC09"
/translation="MNLLEHMGKAQKQASQMLAMLSAKKNOALVIANLLESQTI
LQANERDAAARSGSEALLDRLLITLPARLAIANDVQVCLNDPVGRLDGLSD
SGKLRRRVPVLGVIGVYEARNVTIDVASLCLKTGNVILRGCKETHYINQATNVV
IQALQCGPLFAADVAIESPDRQLVNELRLDRYDMLIPRGASLHKUCREQSTIP
VITGIGCVCHTFDENADFEKALIVENAKIQPSACNSLETLLVHQAVAKTFPLPS
IRHAFGVTLHAGSLAMPYLADGKAKVAVEAADYDDEWLSLDINVIDVTDIAAIDH
IREHGTSHDAIILTRSLSHAEYFVRAVDSSAVVYNASTRFTDGGQFGLGAERAVSTQK
LHARGPMGLDALITYKWIGYGDDLVRS"
complement (4572. .5744)
/gene="proA"
/notes="Pfam match to entry PF00171 aldedh, Aldhyde
dehydrogenase family, score -217.90, E-value 0.00013"
complement (4776. .4841)
/gene="proA"
/notes="PS01223 Gamma-glutamyl phosphate reductase
signature."
```

	Query Match	36.0%;	Score 536.6;	DB 1;	Length 220050;
	Best Local Similarity	69.7%;	Pred. No. 9.5e-148;		
	Matches	740;	Conservative 0;	Mismatches 319;	Indels 2; Gaps 1;
Qy	429	TATTTTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTCTGTTCGATCAATAAGTT	488		
Db	121630	TATCAATAATAGATAGAAATCTATCATATTAATAGGACTAACGAATTTGATTTTCCAAFTT	121689		
Qy	489	GACITTTGTATACACATGCGAATTTCCCTTAATCCGGAGCTATTCGTATGATGATAAAAA	548		
Db	121690	AATCATGTAAAAAATAGTTGCCCTTTAGTTTAAAGGGGACTTAT--ATATGTTAAAAAAA	121747		
Qy	549	ACTCTTCTCTGTCTGATTTCTTCGGCGTATCGGGGAGCTTTTCTACCGCTGTAGCCGCT	608		
Db	121748	ATCTTACCGGTACTTAATACTTCGCCATTGTACATATACACCTACGGCTTGGGCAGCA	121807		
Qy	609	GATAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTCTCTGTAGA	668		
Db	121808	GAAGCACCAGACTGATAGTTTTTACTTACCCAAAAGCCTCGATCTCTCCACTCCGT	121867		
Qy	669	TTACACAGCCCTGAATCAATCTCTGGGGGCTGATTTTGATTTATGCCACACAGATTCAA	728		
Db	121868	TTACATAATATAGAGTCCATCTCTACGGAAGGATTTTAATATATGCTCAACAGTTAAA	121927		
Qy	729	CAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTTGCTGACAACTTTCCAGGAT	788		
Db	121928	ACACTGGATCTTGAAGCAGTAAAGAGATATAAAACAGTCTTACCAATCACAAGAC	121987		
Qy	789	TGGTCCCTCGGATATGTCATATAGTTCCTTTCTTTTATTCGTATGGCTTGGCACCGT	848		
Db	121988	TGGTGGCTGCTGATTATGGTAAATATATGGTCCATTTCTTTTATTCGTATGGCTGGCATG	122047		
Qy	849	CCCGGAACATACAGGACATATGATGGCGGGGAGCGCCAGTGGTGGTCAGCAACGTTTT	908		
Db	122048	CGCGGAACCTACCGCATATATGATGGCGCTGGTGGTGGCGATGGCGGGCAGCAAGATTT	122107		
Qy	909	GAAACCGCTGAACAGCTGGCCGGATAACCGTTAATCTGGATAAAGCCCGCTCGATTCGTGG	968		
Db	122108	GAGCACTCAATAGCTGGCCAGATTAACGCCAACCTTGTATAAGCCGCTCGGCTCTGTGG	122167		
Qy	969	CCAGTCAAGAAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTCTCTGACTGGT	1028		
Db	122168	CCTATAAAAAAGAAATATGGCGCTAAAAATATCGTGGGGCGACCTGATGCTCTTACAGC	122227		
Qy	1029	AATGTTGCCCTTGAATCCATCGGATTTAAACCGCTGGGATTTGCTGGCGGAGAGAAGAT	1088		
Db	122228	AATGTCGCGCTCGAATCTATGGGGTTTAAAAACGCTGGGGGTTTTCAGGAGGCCGCTGAGGAT	122287		

Qy	1089	GACTGGAGTCGGACCTGGTATCTACTGGGGGCTGACAAAGCCCTCTTGCGAGATAACCGG	1148
Db	122288	GATTGGCAATCGATCTGGTCTACTGGGGGCGAGCAACAGATGCTGTCTGATAACAGG	122347
Qy	1149	GATAAAACGGGAACTTTCAGAAACCTCTTGGCGGCACGACGATGGGACTTATTTATGTC	1208
Db	122348	GATAAAATGCAAGCTACCCAAACCGCTGGCGGCACACAAATGGGGTGTGATTTACGTA	122407
Qy	1209	AATCTGAAGCCCGGTCGGAACACCATCTCTGGCTTCGCGGAAAGATATCAGGAA	1269
Db	122408	AACCTGAAGGACCAATGGTAACACGACCGGTTGCTGGCGCAAAAGATATTCGTGA	122467
Qy	1269	GCTTTTTCACGATGCGCATGGATGATGAGGAGCTGTGGCCCTCATCGCGGAGGCGAT	1328
Db	122468	GCGTTTTCACGATGCGCATGATGATGAGGAGCTGTGGCCCTCATCGCGGAGGCGAT	122527
Qy	1329	ACATTTGTAAGCACATGTCGACGCTCTCTGAAAAATGATTTGGCGCAGGCGCTGAT	1388
Db	122528	ACCTTCGTAAGCGCATGGCGCAGCTCTCTGAAAAATGCTGGGTGCAGCACCGGCT	122587
Qy	1389	GGTCACCTGGGAGGACGAGGACTGGATGGGAAAAATAATGTCGTACAGGAAACGGC	1448
Db	122588	GAACGGGGCTAGAACACAAAGGTTTAGGATGGGCGAATAATGTCGTACAGGTAATGCT	122647
Qy	1449	AAATATACCATCACACAGTGGCTGGAAGGAGCCTGGTTCGAC	1489
Db	122648	AGGACACGATACACAGTGGGTTGGAAGGCGATGGACAC	122688

RESULT 12
AE017128/c
LOCUS
DEFINITION
Yersinia pestis biovar Medievalis str. 91001 section 2 of 16 of the complete genome.
ACCESSION
AE017128 AE017042
VERSION
AE017128.1 GI:45434997
KEYWORDS
SOURCE
Yersinia pestis biovar Medievalis str. 91001
ORGANISM
Yersinia pestis biovar Medievalis str. 91001
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Yersinia.
REFERENCE
1 (bases 1 to 290803)
Song, Y., Tong, Z., Wang, L., Han, Y., Zhang, J., Pei, D., Wang, J., Zhou, D., Han, Y., Pang, X., Zhai, J., Chen, F., Qin, H., Wang, J., Li, S., Guo, Z., Ye, C., Du, Z., Lin, W., Wang, J., Yu, J., Yang, H., Wang, J., Huang, P. and Yang, R.
Direct Submission
Submitted (24-APR-2003) The Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, No. 20, Dongdajie Street, Fengtai District, Beijing 100071, People's Republic of China
TITLE
JOURNAL
FEATURES
source
1. 290803
/organism="Yersinia pestis biovar Medievalis str. 91001"
/mol_type="genomic DNA"
/strain="91001"
/db_xref="taxon:229193"
/note="biovar: Medievalis"
complement (314. 1168)
/gene="panC"
/locus tag="YP0283"
complement (314. 1168)
/gene="panC"
/locus tag="YP0283"
/note="highly similar to Yersinia pestis C092 panC and Yersinia pestis KIM panC. Contains Pantoate-beta-alanine ligase (Interpro|IPR003721, (GO:0015940))"
/codon_start=1
/transl_table=11
/product="pantoate--beta-alanine ligase"
/protein_id="AAS60558.1"
/db_xref="GI:45434998"
/translation="MLIETPLLRQRIIRWRQEGKKIALVPTMGNLHEGHMTLVDEA

gene
CDS
gene
CDS
gene
CDS
gene
CDS

KTRADVVTITFVNPLQPERDDLAHYPRTLQEDCEKLTIRHGADLVFAPAAADIYPAG
LEKTYVDVPLSTILEGASRPHGRGVSTIVSKLFNLIQPDVACFGKDYQQLALIR
KMWADMGYDINIVGVTVRADKGLASRNGYLTEERQIAIPLQSKIMWALAEKWLK
ERQIDALLEEAAQAQLRVGTFPDELFIIRDAETLQPLTVDVDSQAQVILMAWLGKARLID
NQLVDLRH"
complement (1256. .2053)
/gene="panB"
/locus_tag="YP0284"
complement (1256. .2053)
/gene="panB"
/locus_tag="YP0284"
/note="highly similar to Yersinia pestis C092 panB and Yersinia pestis KIM panB. Contains Ketopantoate hydroxymethyltransferase (Interpro|IPR003700, (GO:0015940))"
/codon_start=1
/transl_table=11
/product="3-methyl-2-oxobutanoate hydroxymethyltransferase"
/protein_id="AAS60559.1"
/db_xref="GI:45434999"
/translation="MKTITTSQLRQWKQEKRFATLTAYDASFAQLFAEQGIQVLLVG
DSLGMTLQGGDSTLPTVADVMAYHTRAVRGAPHGLLIADMPFMSYATPELAFTAAE
LMRAGANNVKLGSSWLCDDTIRMLAERAVPVCGHLGTPQSVNIFGGYKVGQREEVAA
NQLQDALEQAQAQLLVLECVFVELAQRTESLTIPVIGIGAGNVTDGQILVWMDA
LGIYTGHTPKFSPKFLAHSAGDITRAAIKLYIEVEGGIYPAEHTFQ"
complement (2322. .2801)
/gene="folK"
/locus_tag="YP0285"
complement (2322. .2801)
/gene="folK"
/locus_tag="YP0285"
/note="highly similar to Yersinia pestis C092 folK and Yersinia pestis KIM folK. Contains 7,8-Dihydro-6-hydroxymethylpterin-pyrophosphokinase, HPPK (Interpro|IPR000550, (GO:0009396))"
/codon_start=1
/transl_table=11
/product="2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase"
/protein_id="AAS60560.1"
/db_xref="GI:45435000"
/translation="MIRVYIALGSNLAMPLOQVSAAREALHPRSLRVACSPLYRTK
PLGQDPDFENAVVALDTSPPQLDHTQAIERNQVRKQERWGRPTLDDLDIMLY
GDQVTKDRLTIHYGLKAREFMLYPLADIAPDLIFPDGESLSLSECLKRVKNGVLW"
complement (2809. .4230)
/gene="pcnB"
/locus_tag="YP0286"
complement (2809. .4230)
/gene="pcnB"
/locus_tag="YP0286"
/note="highly similar to Yersinia pestis C092 pcnB and Yersinia pestis KIM pcnB. Contains Polynucleotide adenyllyltransferase (Interpro|IPR002646, (GO:0006396))"
/codon_start=1
/transl_table=11
/product="poly(A) polymerase"
/protein_id="AAS60561.1"
/db_xref="GI:45435001"
/translation="MPCGDNVARKERPAPATGKRGHAVSSSFQROMAITPRDOHNI
SRDISDNPLKVLRLNKSQYEAIVLGGVDRDLLLGRKPDFDITTSATPQVRKLFPR
NCRIVGRRFLAHVMFGFEIIEVATFGHHEQQQAEDSKNSQAQNGMLRDNIFML
SIEDAQRRDFTINSYLYSDFAIRDYTGGLDKGIIIRLIGDPTRYREDVPMFL
YQVFLPAKLDMISIPETAETPIPRLASLIRIPARLFEESKLQSGYGYTKYKLCE
YQLFQPLFIARNFTQHDSPMERILVOVLKNTDRLHNDQRVNPAFLPAALWYPL
IEHAQKLTQSSGLAYDAFALAMNDVDECSRSLAIKRIITSLVDRDWLQRLSRQ
GKAHKLMEHPKFRAYDALLLRAEVKNEHQLQRAWGEFQETATPQQSKMLNTLIG
ADPAPRRSRPRPKVPVRKEG"
complement (4373. .5527)
/gene="glnS"
/locus_tag="YP0287"
complement (4373. .5527)
/gene="glnS"

```
/locus_tag="yp0287"
/note="highly similar to Yersinia pestis CO92 YP03398 and
Yersinia pestis KIM y0789. Contains Glutamy1-tRNA
synthetase, class Ic (Interpro|IPR000924, (GO:0006424))"
/codon_start=1
/transl_table=11
/product="putative glutamyl-tRNA synthetase"
/protein_id="AAS60562.1"
/db_xref="GI:45435002"
/translation="MLEDTGFQRLFSQLPDPQPFQSGKPSQLFGSQPFAAQLFSA
AOPRAGFPQAOLPETQPVQAOVQIORSANQATNPTEYIGRFAPSPSGDLHFGSL
IAALGSVLOARAGKWLVRIEDIDPREVPGAASRILAALEHYGLHWDGPIYQSOR
HEAYRATLWLEQGLSYCTCRSRHQLGGFVGVCDRHLHPASGAAIRLQTOPV
YAYDKLGLBHAHPALAAQEDFIIRRDGGLFAYNLAVVDADPQGVTEIVRGADLIEP
TQROIALYQOOLHPFVGYIHPLALNNQGNKLSQNHAPPELDPAPILIDALKELR
QPLPEYQWQDLDLYLLRYAYEVRHTLVSIPLOQGAITPQKTRHSQSKYGEI"
5419..5448
/note="degenerate repeat region containing 2 copies of a
15bp unit"
/rpt_type=tandem
5579..5612
/note="degenerate repeat region containing 2 copies of a
17bp unit"
/rpt_type=tandem
complement(5618..6166)
/gene="dksA1"
/locus_tag="yp0288"
complement(5618..6166)
/gene="dksA1"
/locus_tag="yp0288"
/note="highly similar to Yersinia pestis CO92 dksA and
Yersinia pestis KIM dksA. Contains Zn-finger, prokaryotic
DksA/Trar C4 type (Interpro|IPR000962)"
/codon_start=1
/transl_table=11
/product="DnaK suppressor protein homologue"
/protein_id="AAS60563.1"
/db_xref="GI:45435003"
/translation="WPYLFVSIVYSGLIFFPIGDFDAYSRECRSIMQEGQKRTSSLS
ILAIAGVEPYQEKPEGYMNAQLSHFKLILEARNQLRDEVDTRVSHMQDEAANFPD
PVRAQOEBSFLELRNRDRERLKKIEKTKKVEDDDFGFCSCGVGEGIRRLR
PTADLCIDCKTLABIREKQWAG"
complement(6243..7118)
/gene="sfaA"
/locus_tag="yp0289"
complement(6243..7118)
/gene="sfaA"
/locus_tag="yp0289"
/note="highly similar to Yersinia pestis CO92 sfaA and
Yersinia pestis KIM sfaA. Contains Sugar fermentation
stimulation protein (Interpro|IPR005224)"
/codon_start=1
/transl_table=11
/product="sugar fermentation stimulation protein"
/protein_id="AAS60564.1"
/db_xref="GI:45435004"
/translation="MKNQMPPANANLPPFSANLVSNSDSRSISPNPPTHAKSLQFN
APLIQFTPPLOPALTILIRYKRLADIIVTPAGEALTIHCANTGAMTGCATPGDITWIST
SNDPKRYPQSWELTQTQTDWICVNTWRANELVNLAIERNQIAELSGSNFVRKEVY
GEENSRIDLLQAQRDCYIEVKSATLLQOCCGVPDAVTLRGQKHLRELQNRVWG
HRAVLPFNAVLTGILKQVAPARHIDRYAELLVQAQAGVEICVGFQLSPDGIELNTR
LPLJDEMLSSENAE"
complement(7093..7668)
/gene="ligT"
/locus_tag="yp0290"
complement(7093..7668)
/gene="ligT"
/locus_tag="yp0290"
/note="highly similar to Yersinia pestis CO92 ligT and
Yersinia pestis KIM y0793. Contains 2_5' RNA ligase
(Interpro|IPR004175, (GO:0016070))"
/codon_start=1
/transl_table=11
```

```
/product="putative 2'-5' RNA ligase"
/protein_id="AAS60565.1"
/db_xref="GI:45435005"
/translation="MTNTVDELKRVRLFPALTLPDAMQOEIIVOMRAGHFSPEAGRPV
AAANLHLTLAFLEVSATSKSOVLQOAGRIQOAGFVTLDDIGHWPGSGVILGCKNP
PRGLLQALLRSQAARSQCYQTPFLPFHPHVTLIRNATRPVALPAKSGNETFQADHFS
LYESVPARGTRYNIYQSWPLAGSERKPDAC"
7709..10270
/gene="hrpB"
/locus_tag="yp0291"
7709..10270
/gene="hrpB"

Query Match 36.0%; Score 536.6; DB 1; Length 290803;
Best Local Similarity 69.7%; Pred. No. 1e-147;
Matches 740; Conservative 0; Mismatches 319; Indels 2; Gaps 1;

Qy 429 TATTTTTTGTCTGTAGTATCTATCCAGCAATAGTATATCTCTGTTCGATCAATAAGTT 488
Db 98936 TATCAATAATAGATAGAATCTATCATATTAATAGGACTAACGAATTTGTTTCCAAATTT 98877
Qy 489 GACTTTTGTATACAACATGCGAATTTCCCTTAATCCGAGCTATTCTGATGATATAAAAA 548
Db 98876 AATCATGTTTAAAAAATAGTTGCCCTTTTAGTTAAAGGGGACTTAT--ATAIGTTAAAA 98819
Qy 549 ACTCTTCCTGTTCTGATTCTTCTGGCGCTATCGGGGAGCTTTTCTACCGCTGTAGCCGCT 608
Db 98818 ATCTTACCCGTACTAATACTCTCGCCATTGTACATATAACACCTACGGCTTGGGCAGCA 98759
Qy 609 GATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTTAACCTCTCGAGA 668
Db 98758 GAAGCACCAGACTGATAGTTTTTACTTACCCAAAAGCCCTCGATCTCTCTCCACTCCGT 98699
Qy 669 TTACACAGCCCTGAATCAATCCTCGGGGGCTGATTTTGATATATGCCACAGATTTCAA 728
Db 98698 TTACATAATATAGAGTCCAATCCTTACGGAAGAGATTTTAATATGTCACAGTTTAA 98639
Qy 729 CAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTTGTCTGACAACCTTCCCAGAT 788
Db 98638 ACCTGGATCTTGAAGCAGTAAGAAGATATAAAACAGTCTTACCACATCAACAGAC 98579
Qy 789 TGGTGCCCTCGGATATATGGTCATATATGGTCTCTTTCTTTATTCGTATGGCTTGCACCGT 848
Db 98578 TGGTGGCTGCTGATTTATGGTAAATATATGGTCCATTTCTTTATTCGTATGGCTGGCATGGT 98519
Qy 849 GCCGGAACATACAGACATATGATGGCGGGGAGCGCGCTGCTGGTGCAGCAACGTTTT 908
Db 98518 GCGGGAACCTACCCGATATATGATGGCGCTGGTGGTCCGATGGCGGCAGCAAGATTT 98459
Qy 909 GAACCGCTGAACAGCTGGCCGGAATAAGTAACTGGATAAAGCCCGCTCGATTGCTGTGG 968
Db 98458 GAGCCACTCAATAGCTGGCAGATAACGCCAACCTTGATAAAGCGCGTCCGGCTTCTGTGG 98399
Qy 969 CCACTCAAGAAAAAATACGGCTCCAGTATTTCTCGGGGAGACCTGATGCTCTGATCGT 1028
Db 98398 CCTATAAAAAAGAAATATGGCGCTAAAAATATCGTGGGCGACCTGATGCTCTTACAGGC 98339
Qy 1029 AATGTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAAGAT 1088
Db 98338 AATGTCGCGCTCGAATCTATGGGGTTTAAACGCTGGGGTTTGCAGGAGGCCGTGAGGAT 98279
Qy 1089 GACTGGGAGTCGGACCTGGTATATCTGGGGGCTTGACAAACAAGCCCTCTTTCAGATACCCGG 1148
Db 98278 GATTGGCATCCGATCTGCTCTACTGGGGGCGAGCAACAAGATGCTGCTGTATAACAGG 98219
Qy 1149 GATAAAAAAGGAAATCTTCAGAAACCTCTTGGCGGCACGAGATGGGACTTATTTATGTC 1208
Db 98218 GATAAAATGCAAGCTACCCAAAACCGCTGGCGGCAGACACAAATGGGGTTGATTTACGTA 98159
Qy 1209 AATCCTGAAGCCCGGCGGAAACACAGATCTCTGGGCTTCGCGGAAAGATATCAGGGAA 1268
Db 98158 AACCTGAAGGACCAANAATGGTAAACCGACCCGGTGTCTGCGGCAAAAGATATTCGTGAA 98099
```



```
/translation="MAERRIPLKHEVAQDKVIYGVDEAGRGFLAGPVVAAACVVLQPA
DVTGLADSKQLSEKRIISLADQIKQARAWAJASAVEBIDRLNIIQLASLLAMQRAY
VSLRPSNALVLDVGDHAPRLDCEVOTIRGDSLVAEISAAITAKTARDIEMLRHE
AVPVGDRHKGYPTKAHLRAIRLHGITDIHRRSFAPCVGQSVSGARTTSFINOKEA"
complement (5068. .5529)
/gene="fabz"
/locus_tag="NE1708"
complement (5068. .5529)
/gene="fabz"
/locus_tag="NE1708"
/EC_number="4.2.1.-"
/function="InterPro IPR001143 COGs COG0764"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="Bacterial thioester dehydrase"
/protein_id="CAD85619.1"
/db_xref="GI:30180707"
/translation="MKQGDQRTVMDTHEILKYLPHRYPLILVDRVVSLESGKRIHAVK
NFSINEPYFGHPHHPVMPGVLLIEALAAQAAALLTIRSENMEKONGQVYFVGDIV
RFKKPVTAGDLVLKVEITRLQKGIWKYAAAEVDGQVVTEAQLMCTARAI"
complement (5546. .6109)
/locus_tag="NE1709"
complement (5546. .6109)
/locus_tag="NE1709"
/function="COGs COG2825"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative transmembrane protein"
/protein_id="CAD85620.1"
/db_xref="GI:30180708"
/translation="MRYPLDQYMQMFFRFVKAFFVVMFVLPHVSSAGBIKIGVNT
EKLVRSEMPAIEAQKIEREFOARDARIELSAQITALQOELEKNTGTVDDEERLKE
RELAGLSQVORQAQQWREDLSLQNEEYGLILERINQVIRELAEKQSYDLILQLQDS
VYRSARIDIDQVILKVNARESAKRP"
complement (6005. .6070)
/locus_tag="NE1709"
/note="1 probable transmembrane helix predicted by
TMHMM2.0"
complement (6136. .8412)
/locus_tag="NE1710"
complement (6136. .8412)
/function="InterPro IPR000184 COGs COG0729"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="Bacterial surface antigen (D15)"
/protein_id="CAD85621.1"
/db_xref="GI:30180709"
/translation="WKLRFLILFSLYSLGCMANDSLVVRDIRVEGIQRTBAGTVFSY
LPVKGVLDSSKSAIKALYATGFFSDVKLKEGGLLI VOVERPAIAQISINGAK
EFDKLEKGLKQAGLSESRIFPSLLEKAEOLKROVISRKYAVKLTITTTPLERN
RIGFIDLEKGTARIKQINIVGNHVPFEDDLVLFLSLKTPGFWTFKDDQYSKQKL
SADLETLSRYLDLGFLEFNTEQVSLTPDMKDIYITVNEGQPYTVSDIKLAGE
LVPEERLKLKLFEGGIFVREKLTESKLJSDRLGNDGYAFANVPASPELDKTRKT
AFTFIDPGRVYVRDINISGNERTDRVIREFRQMGWHSTEQINRSQRVDRLD
PFTGVNIEPTPVADVPQVDINNVVEKPTGAIMFGAGYSDRGIIILNGSIAQNNILG
TGNFLSQVNTGSKVNTSASPNPYTYTINGVSLGLEAFKRDINTRSLSSVMFNTDT
TGANRIFGIPVAENDIVSLGLGYEHTKIDLRDSDSPQRFKDFVDFQKTSNNLPITLS
ARDNRNSAIWTTSGMNTQRLFGPLFGDLNLYKYVSEQRWFFPVTKMTFLMNGEVS
VGDGYSKPLPFPFNKFGAGFNVRGYNINTLIPRDSDDRLVGLGSKRIVGNIEVLPV
PFMKEDKSVLSAFADGGTIVNSFSGLGFDFFRYSAGLAATVSPMGPLKFSVAQPLN
NQSGLKLFQFQFGTF"
complement (8415. .9782)
/locus_tag="NE1710"
/note="Signal predicted by SignalP 2.0 HMM (Signal peptide
probability 1.000) with cleavage site probability 0.998 at
residue 20"
complement (8415. .9782)
/locus_tag="NE1711"
complement (8415. .9782)
```

```
Query Match 26.5%; Score 394.6; DB 1; Length 295250;
Best Local Similarity 66.8%; Pred. No. 2.6e-105; Gaps 1;
Matches 581; Conservative 0; Mismatches 289; Indels 3;
QY 620 GACTCAAAATTTCTACTATCCAGAAACACATCGGATTTTAACTCTCTGAGATTACACAGCCC 679
DB 125602 GATGAACAGTTTTTGGTGGCCGGATCAACTGGATCTGAGACCGTTGCGCCAGAATTCGGC 125453
QY 680 TGAATCAAAATCCCTGGGGGGCTGATTTTGTATTATGCCACAGATTTCACAGCTGGGATAT 739
DB 125542 TGAGTCTAACCCCTTGGGCAAGCGTTTCAACTATATGCCGAAACAGTTTTTAAACGCTCAACCT 125483
QY 740 GGAGGCTCTGAAAAGAGATATCAAGATTGTCTGACAACTTCCACGAGATTGGTCCCTGTC 799
DB 125482 CAAAGCGGTGAAGAGATATCGGAAAGCTGCTGCATATCTCCAGCTTGGTGGCCGGC 125423
QY 800 GGATTATGGTCAATTATGGTCTTTTCTTTATTCGTATGGCTTGGCACCGGTGGCGGAAACATA 859
DB 125422 GGACTAGGCAACTACGGACCAATTGTTCAITTCGCATGGCTGGCATAGTGGCGCGTGA 125363
QY 860 CAGACATATGATGGCCGGGAGCGCCAGTGGTGGTTCAGCAAGCTTTTGAACCGCTCAA 919
DB 125362 CCGCAITTTTCAGCGAACGTGTGGCGCTTCCGGCGGATAGCAACGCTTCGAAACCGCTCAA 125303
QY 920 CAGCTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCGATTGTCTGTGGCCAGTCAAGAA 979
DB 125302 CAGTTGGCCGGATAACGTTAATCTGGATAAAGCCCGTCGATTGTCTGTGGCCAGTCAAGAA 125243
QY 980 AAAATACGGTCCAGTATTTTCTGGGGAGACCTGATGCTCCTGACTGCTGATTAATGTGGCCT 1039
DB 125242 GAAATACGGCAGCAAGCTTTCTTGGGCTGATCTGATGGTGGCGCAATGTGGCGTT 125183
QY 1040 TGAATCCATCGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAAAGATGACTGGGAGTC 1099
DB 125182 GGAGGATATGGGATTTAAACCATAGTGTGGCGGTGGCGGACCGGATGATGGGAAGC 125123
QY 1100 GGACCTGGTATATCGGGGGCTGACAAACAGCCCTCTTGCAGATAACCGG---GATAAAA 1156
DB 125122 CGAGATCGTCAACTGGGGAAACAGAAAAGGAATTTCTCGCGATGAGCGCCATGATAAAG 125063
QY 1157 CGGAAAACCTTCAAGAAACCTCTTTCGCCGACGACAGATGGGACTTATTTATGTCATCTGA 1216
DB 125062 AGGTGAATGGCGAAGCTCTGGCTGCGGTGCAGATGGGGCTGATCTACGTCACCCGGA 125003
QY 1217 AGGCCCCGTGGAAAAACAGATCTCTGGCTTCGCGGAAAGATATCAGGGAAGCTTTTC 1276
DB 125002 AGGCGCGGGCGGTAAACCCGATCCGTTGGCAGCTGCCAGGCATATCCGCGAATCCTTCGG 124943
QY 1277 ACGTATGCCATGATGATGAGGAGACTGTGGCCCTGATCCGCGAGGCGCATACATTTGG 1336
DB 124942 CCGTATGGCGATGAATGACGAAGAGACGCTGGCTCTGATCCCGGTGGCCATACCTTCGG 124883
QY 1337 TAAAGCATGGTGCAGCGTCTCTCTGAAAAATGTTATGGCGCAGGCGCTGATGTCACCC 1396
DB 124882 CAAGCGCACGGGGCGCATAGCCGGAAGATGCGTTGGCAAGGAACCGGCGCGGCTGG 124823
QY 1397 TGTGGAGAGCAGGGAATGGGATGGAAAAATAAATGTTGGTACAGAAACGCAATATATAC 1456
DB 124822 TATCGAGAGCAAGGTTTGGGTTGGGCGCAATATAATGTTGGCAGCGGCTATGGCGTCGATAC 124763
QY 1457 CATCACCAGTGGCTTGGAAAGGAGCGTGGTCGAC 1489
DB 124762 AGTCACCAGTGGTCTTGGAAAGGCGGCGTGGTCGAC 124730
```

RESULT 14
AB017595
LOCUS
DEFINITION
AB017595
AB017595
AB017595.1
GI:4996127
katA; catalase-peroxidase.

AB017595
Legionella pneumophila gene for catalase-peroxidase, complete cds.
AB017595
AB017595.1
GI:4996127
katA; catalase-peroxidase.

SOURCE	Legionella pneumophila	Qy	890	TGCTGGTACCAACGCTTTTGAACCGCTGAACAGCTGGCCGGATAACGTTAATCTGGATAA	949
ORGANISM	Legionella pneumophila	Db	571	CGGTGGCTTCAGCGTTTTCGCCCGCAAAACAGCTGGCCGGATAATCCCAATCTGGACAA	630
REFERENCE	1 (sites)	Qy	950	AGCCCGTGCAGTTGCTGGCCAGTCAAGAAAAAATACGGCTCCAGTAGTATTTCTCTGGGAGA	1009
AUTHORS	Anemura-Maekawa,J., Mishima-Abe,S., Kura,F., Takahashi,T. and Watanabe,H.	Db	631	AGCAGACGTTTACTGTGGCCAAATTAAGCAGAAATATGGCGGCAGAGATTTTCATGGGCTGA	690
TITLE	Identification of a novel periplasmic catalase-peroxidase KatA of Legionella pneumophila	Qy	1010	CCTGATGTCCTCACTGATTAATGTTCCTTGAATCCATCGGATTTTAAACGCTGGGATT	1069
JOURNAL	FEMS Microbiol. Lett. 176 (2), 339-344 (1999)	Db	691	TTTGTGGTACTTGTGGAATGTTCGATGGAATCTATGGGCTTCAAGACCATTTGATTT	750
MEDLINE	99356723	Qy	1070	TGCTGGCGGAAGAGACATGACTGGGAGTCGGACCTGGTATATCTGGGGCTTCGACACAA	1129
PUBMED	10427716	Db	751	TGCTGGAGGCGGTGAAGACGATGGGAGGCCATCAATATCAATTTGGGGCCCTTGAAGGCAA	810
REFERENCE	2 (bases 1 to 2587)	Qy	1130	GCCTCTTTCAGATAA---CCGGGATAAAAACGGGAAACTTTCAGAAACCTCTTTCGCGCAC	1186
AUTHORS	Anemura-Maekawa,J. and Watanabe,H.	Db	811	ATGGCTGGAAGTAAGCGCCAGGATAAGATGGAANAACCTTGAAAAACCGCTTGTCTGGAC	870
TITLE	Direct Submission	Qy	1187	GCAGATGGGACTTATTTATGTCAATCTGTAAGCCCGGTGGGAAACAGATCTCTTGGC	1246
JOURNAL	Submitted (10-SEP-1998) Junko Anemura-Maekawa, National Institute of Infectious Diseases, Department of Bacteriology, Toyama 1-23-1, Sinjuku-Ku, Tokyo 162-8640, Japan (E-mail:jmaekawa@nih.go.jp, Tel:81-3-5285-1111, Fax:81-3-5285-1163)	Db	871	CGTAATGGGCTTAATCTATGTGAATCGGAAGCACAACCGCGTTCCAGATCCTCTTGC	930
FEATURES	Location/Qualifiers	Qy	1247	TTCCGGGAAGATATCAGGGAAGCTTTTTCAGTATGGCCATGATGATGAGGAGACTGT	1306
source	1. 2587	Db	931	TGCGGCGCAAAAAATTTTCGGGAGACTTTTCGGGCGGTATGGCCATGAATGACGAAGAACAGT	990
gene	/organism="Legionella pneumophila"	Qy	1307	GGCCCTGATCGCGGGAGGCATACATTTGTTAAAGCACATGGTGCAGCGCTCTCTCGAAAA	1366
CDS	/mol_type="genomic DNA"	Db	991	TGCTTAAATTCGCGGGCGGCACCGATTCGAAAAACACATGGAGCCCGT---CCGCAAA	1047
	/strain="AM511"	Qy	1367	ATGATATTGGCGCGAGGCGCTGATGTCACCTGTGGAGGACGAGGACTGGGATGGAAAAA	1426
	/db_xref="taxon:446"	Db	1048	ATATCTGGGCGCAGCGCCAGCAGCAGCATAGAAGAACAGGCTTTGTTGGAAAAA	1107
	203..2452	Qy	1427	TAAATGTGTACAGGAAACCGGCAATATACCATCACCAGTGGCTGGGAAGAGCTGCTC	1486
	/gene="katA"	Db	1108	CAGTTACGCGAGCGGCAAGGGAAGACACGATTACCACTGGTGTGTTAGAAAGTGCCTGGAC	1167
	/codon_start=1	RESULT 15			
	/transl_table=1	AP276752			
	/product="catalase-peroxidase"	LOCUS			
	/protein_id="BA078342.1"	DEFINITION			
	/db_xref="GI:4996128"	AP276752			
	203..2452	AP276752.1			
	/gene="katA"	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	/gene="katA"	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			
	203..2452	AP276752			
	203..2452	AP276752.1			
	203..2452	GI:11528084			
	203..2452	LEGIONELLA PNEUMOPHILA			
	203..2452	5249 bp DNA linear			
	203..2452	Legionella pneumophila catalase-peroxidase (katA) gene, complete cds.			

/notes="compared to Legionella pneumophila kata sequence in
GenBank Accession Number AB017595"
/replace="g"

misc_difference

2227 /notes="compared to Legionella pneumophila kata sequence in
GenBank Accession Number AB017595"
/replace="c"

gene

2386. .4635 /gene="kata"
2386. .4635 /notes="kata"

CDS

/notes="kata; localized in periplasm"
/codon start=1
/transl table=11

/product="catalase-peroxidase"

/protein_id="AAG37106.1"

/db_xref="GI:11528085"

/translations="MFKRTIPLFAAFTLAIISPFIIPNYAHQAEDKPTNQYWPVKMLD
LSRPNATSNPMGKFNABEFNSLDLNAVIEDLKLMTTSQDWWPADIYNGYPLF

IRMSWAAGTYRIYDGRGGANGGFORPAPONSFPDANLDKARLLWPIKOKYGRKIS
WADLLVLAGNVAMESGFKTIGPAGGREDWEALINWPEGKWLSEKQDKDGKLEK

PLAATVWGLIYVPEGGNGVDFPLAAAKIRETFGRMANDEETVALIAGGHAFKTH
GAASGYLGPAPAGAEQEGQKNSYSGSGKDTITSGLEGATVTPTHWSHNYLQ

NLFNFVWTKSPGGAIQWFPENSASVDPDAFDPKSRHAPVMTLTDLAKFPDYPYS
KIAKFLDNPKEDDAFARAWFKLIIRDMPGRSRYLGLSVPEKAMTWQDPVPVDYKL

VDANDIANLKGILNSLTTSVLKTAWASATFRGTDMEGGANGARIILAPQDWPA
NDQBELAKVLKLTLESTQNFNFNAQDGKILSLADLIVLGNAALEQNAKQAGYDIIVP

FTPGRTDITQGMIDTVSFEVLEKADGRFNYFDKSNMSPPEMLVEKASLLKUSVEM
TVLVGGMRVNAVNTQNGYVFTDKPTGLNNDFFINLLMSNTWKSSSETEGIYEGYE

RKTGKLKATSVDLIFGANSERLVAEAYATDDAKEKFTIQDFINAWVKVMTADRFDI
KAANANINS"

ORIGIN

Query Match 25.8%; Score 384.8; DB 1; Length 5249;

Best Local Similarity 65.8%; Pred. No. 8.3e-103;

Matches 592; Conservative 0; Mismatches 302; Indels 6; Gaps 2;

Qy 590 TTCTACCGCTGTAGCCGCTGATAAAGAGACCTCAAAATTTCTACTATCCAGAAACACT 649

Db 2454 TTATGCCCATCGGCAAGACAGCCGAAACCAATCAATACCTACTGTGGCTTAAATGCT 2513

Qy 650 GGATTTAACTCCTCTGAGATTACACAGCCCTGTAATCAATCCCTCGGGGGCTGATTTGA 709

Db 2514 CGATCTTAGCCGCTGCGCAACCAATGCCACTTCGAATCCCATGGGTGAAAAATTCAA 2573

Qy 710 TTATGCCACAGATTTCAACAGCTGGATATAGGAGCTCTGAAAAAGATATCAAGATTT 769

Db 2574 CTATGCCGAAGAATTCATAGCTTAGATTTGAATGCGCTGTATAGAGATCTCAAAAAAT 2633

Qy 770 GCTGACAACTCCAGATTTGGTCCCTCGGATTTATGCTCATTTATGCTCTTTTAT 829

Db 2634 AATGACTACCTCGCAAGACTGGTGGCTGCTGATATGTTAACTATGGTCAATTTATCAT 2693

Qy 830 TCGTATGGCTTGGCACCGTGGCGGAACATACAGACATATGATGGCGGGGAGCGCCAG 889

Db 2694 ACGATGTCATGGATCGGCAGAACTTACCGCATCTATGATGCGCGTGGTGAGCAA 2753

Qy 890 TGGTGGTCAGCAACGTTTTGAACCGCTGAACAGCTGGCCGGATACGTTAATCTGGATAA 949

Db 2754 CGGTGGCTTCAGCGTTTCGCCCGGCAAAACAGCTGGCCGGATATGCCAATCTGGACAA 2813

Qy 950 AGCCCGTCGATGCTGTGGCCAGTCAGAAAAATACGGCTCCAGTATTTCTCTGGGAGA 1009

Db 2814 AGCCAGACGTTTACTGTGGCCAAATTAAGCGAAAAATATGGGCGCAAGATTTTCATGGGCTGA 2873

Qy 1010 CCTGATGCTCTGACTGATATGTTGGCTTTGAATCCATGGGATTTAAAAACGCTGGGATT 1069

Db 2874 TTTGTTGTTACTTCTGMAATGTTGGATGGATCTATGGCTTCAAGACCATTTGATT 2933

Qy 1070 TGCTGGCGGAGAGAAAGATGCTGGGAGTCGGAGCTCGGATCTCGGTATCTCTGGGGGCTTGACAA 1129

Db 2934 TGCTGGAGGGCGTGAAGACGATGGGAGGCCATCAATATCAATTTGGGGCCCTGAAGGCAA 2993

Qy 1130 GCCTCTTTCAGATAA---CCGGGATAAAAACGGGAAACCTTCAGAAACCTCTTTCGCCGCAC 1186

Db 2994 ATGGCTGGAAGTAGTAAGCGCCAGGATAAGATAAGTGAATACTTGAATAACCGCTTGTCTGCGAC 3053

Qy 1187 GCAGATGGGACTTATTATATGTCATCTGAAGCCCGGTGGAAAAACCCAGATCCTCTTGGC 1246

Db 3054 CGTAATGGGCTTAAATCTATGTGAATCCGGAAGGACCAAAACGGCGTTCACGATCCTCTTGC 3113

Qy 1247 TTCGCGAAAGATATCAGGGAAGCTTTTTCACGTATGGCCATCGATGATGAGGAGACTGT 1306

Db 3114 TCGGCGCGAAAAAATTCGCGAGACTTTCGGGCGTATGGCCATGAATGACGAAGAACAGT 3173

Qy 1307 GGCCTGATCGCGGAGGCGATACATTTGGTAAAGCACATGGTGCAGCGCTCTCTGAAAA 1366

Db 3174 TGCTTTAATTGGCGGCGGCGACGCAATTCGAAAAACACATGGAGCGCGT---CCGGCAA 3230

Qy 1367 ATGTATTGGCGAGGCGCTGATGTCACCTGTGGAGGACGAGGACTGGGATGGAAAA 1426

Db 3231 ATATCTGGGCGCCAGCCGCAAGACGACGACATAGAAAGCAAAAGGCTTTTGGTTGGAAAA 3290

Qy 1427 TAAATGTGTACAGGAAACGGCAAAATATACCATCACAGTGGCGCTGGAAGAGCCTGGTC 1486

Db 3291 CAGTTAGGCGCGCGCAAGGGAAGACACGATTACAGTGGTTTAGAAGGTGCTTGAC 3350

RESULT 16

AE017354.02/c

WPCOMMENT

Sequence split into 34 fragments LOCUS AE017354 Accession AE017354

Fragment Name Begin End

AE017354_00 1 110000

AE017354_01 100001 210000

AE017354_02 200001 310000

AE017354_03 300001 410000

AE017354_04 400001 510000

AE017354_05 500001 610000

AE017354_06 600001 710000

AE017354_07 700001 810000

AE017354_08 800001 910000

AE017354_09 900001 1010000

AE017354_10 1000001 1110000

AE017354_11 1100001 1210000

AE017354_12 1200001 1310000

AE017354_13 1300001 1410000

AE017354_14 1400001 1510000

AE017354_15 1500001 1610000

AE017354_16 1600001 1710000

AE017354_17 1700001 1810000

AE017354_18 1800001 1910000

AE017354_19 1900001 2010000

AE017354_20 2000001 2110000

AE017354_21 2100001 2210000

AE017354_22 2200001 2310000

AE017354_23 2300001 2410000

AE017354_24 2400001 2510000

AE017354_25 2500001 2610000

AE017354_26 2600001 2710000

AE017354_27 2700001 2810000

AE017354_28 2800001 2910000

AE017354_29 2900001 3010000

AE017354_30 3000001 3110000

AE017354_31 3100001 3210000

AE017354_32 3200001 3310000

AE017354_33 3300001 3397754

Continuation (3 of 34) of AE017354 from base 200001 (AE017354 Legionella pneumophila sub

Query Match 25.8%; Score 384.8; DB 1; Length 110000;

Best Local Similarity 65.8%; Pred. No. 1.17e-102;

Matches 592; Conservative 0; Mismatches 302; Indels 6; Gaps 2;

Qy 590 TTCTACCGCTGTAGCCGCTGATAAAGAGACTTAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACT 649

Db 25628 TTATGCCCATCGCAAGAGACGCGGAAACCAATCAATACTGTGGTGGCGCTAAAAATGCT 25569

Db	80203	CGTAATGGGCTTAATCTATGTGAATCCGGAAGGACCAACCGCGTTCCAGATCCTCTTGC	80144	gene	/translation="MHNKEKIVVVFIDANKASDKKLDKMLKPIVN"
Qy	1247	TTCCGCGAAGATATACGGAAGCTTTTTCAGTATGGCCATGATGATGAGAGACTGT	1306	CDS	/complement (632..889)
Db	80143	TGCGGCGCAAAAAATTCGCGAGACTTTTCGGGCGGTATGGCCATGAATACGAAGAAACAGT	80084		/locus_tag="SO0712"
Qy	1307	GGCCCTGATCGCGGAGGCGATCATTTGTTAAAGCACATGGTGCAGCGTCTCTGAAAA	1366		/notes="identified by Glimmer2; putative"
Db	80083	TGCTTTAAATTCGCGGCGGCGACGCAATTCGGAAGAACACATGGAGCCCGGT---	80027		/codon_start=1
Qy	1367	ATGATTGGCGCGAGGCGCTGATGTGCACCTGTGGAGGAGCAGGACTGGGATGGAAGAAA	1426		/transl_table=11
Db	80026	ATATCTGGGCGCAGCACCGGAGCAGCGGCATAGAGAACAGGCTTTGTGGAAAAA	79967		/product="hypothetical protein"
Qy	1427	TAAATGTGTACAGGAACCGCAAAATATACCATCACAGTGGCTTGAAGAGGCTGCTGTC	1486		/protein_id="AAN53790.1"
Db	79967	CAGTTACGCGAGCGCGCAAGGAAGATACGATTACCAGTGGTTTAGAAGGTGCTGGAC	79907		/db_xref="GI:24346242"
RESULT 18					/translation="MCCLVASAMGGSVLPSPDNATITRLTPNQKFVKEHKCKMRSRI
AE015517	11421 bp	DNA	linear	BCT 02-DEC-2002	FMIANIVQGEIVPMERITVRVNAVQNSTKYQGEDEINLD"
LOCUS	Shewanella oneidensis MR-1 section 66 of 457 of the complete genome.				
DEFINITION	Shewanella oneidensis MR-1				
ACCESSION	AE015517 AE014299				
VERSION	AE015517.1 GI:24346240				
KEYWORDS					
SOURCE	Shewanella oneidensis MR-1				
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Shewanellaceae; Shewanella.				
REFERENCE	1 (bases 1 to 11421)				
AUTHORS	Heidelberg, J., Paulsen, I., Nelson, K., Gaidos, B., Nelson, W., Read, T., Eisen, J., Seshadri, R., Ward, N., Methe, B., Clayton, R., Meyer, T., Teapin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S., DeBoy, R., Dodson, R., Durkin, A., Haft, D., Kolonay, J., Madupu, R., Peterson, J., Umayam, L., White, O., Wolf, A., Vamathevan, J., Weidman, J., Imbram, M., Lee, K., Berry, K., Lee, C., Mueller, J., Khouri, H., Gill, J., Utterback, T., McDonald, L., Feldblyum, T., Smith, H., Venter, J., Nealon, K. and Fraser, C.				
TITLE	Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis				
JOURNAL	Nat. Biotechnol. 20 (11), 1118-1123 (2002)				
MEDLINE	22297686				
PUBMED	12368813				
REFERENCE	2 (bases 1 to 11421)				
AUTHORS	Heidelberg, J., Paulsen, I., Nelson, K., Gaidos, B., Nelson, W., Read, T., Eisen, J., Seshadri, R., Ward, N., Methe, B., Clayton, R., Meyer, T., Teapin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S., DeBoy, R., Dodson, R., Durkin, A., Haft, D., Kolonay, J., Madupu, R., Peterson, J., Umayam, L., White, O., Wolf, A., Vamathevan, J., Weidman, J., Imbram, M., Lee, K., Berry, K., Lee, C., Mueller, J., Khouri, H., Gill, J., Utterback, T., McDonald, L., Feldblyum, T., Smith, H., Venter, J., Nealon, K. and Fraser, C.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-SEP-2002) The Institute for Genomic Research, 9712				
FEATURES	Medical Center Dr., Rockville, MD 20850, USA				
source	1. 11421				
gene	/organism="Shewanella oneidensis MR-1"				
CDS	/mol_type="genomic DNA"				
	/strain="MR-1"				
	/db_xref="taxon:211586"				
	complement (286..381)				
	/locus_tag="SO0711"				
	complement (286..381)				
	/locus_tag="SO0711"				
	/note="identified by Glimmer2; putative"				
	/codon_start=1				
	/transl_table=11				
	/product="hypothetical protein"				
	/protein_id="AAN53789.1"				
	/db_xref="GI:24346241"				
	/locus_tag="SO0715"				
	/note="identified by match to PFAM protein family HMM PF00174"				
	/codon_start=1				
	/transl_table=11				
	/product="oxidoreductase, molybdopterin-binding"				
	/protein_id="AAN53793.1"				
	/db_xref="GI:24346245"				
	/translations="MDLNRRVFIKGAVATSVLSALPVKWFADKLPAGVSAIDISALT				
	WTKAEDLPSYFKVLTNPLNAPPEHKLAPAVTADVSFIRWNGIVDPDKTLDAKTWT				
	FNYDGSVITPKIYTIIDELKRPKHVTLOLLIRCGNSRAGFPSTKGNWNNSAVFC				
	SEWGLVSDVLDKCGKADAVTGHGADVHLISGKEALSRGVPIKAALBELGLIAM				
	SMNGEDIPYLHGPIRLIVFGGRFSGVSQKATGLISVRNKVHDGKMEAPAYQVPKYPI				
	APGETVDENDYRIIEEMIVKSLITPQSGTLAGQKLTAVGHAWAGRVVKEVSY				
	DYGRWFVAELQPKNATAQQWMSIELQLPTQGYEITWAKATDSEGSQPVVQPMNP				
	KGYMENGCHRIAVRVV"				
	3493..3876				
	/locus_tag="SO0716"				
	3493..3876				
	/locus_tag="SO0716"				
	/note="similar to GP:6967509; identified by sequence				
	similarity; putative"				
	/codon_start=1				
	/transl_table=11				
	/product="monoheme cytochrome c, putative"				
	/protein_id="AAN53794.1"				
	/db_xref="GI:24346246"				
	/translations="MKAHVNKLMCGLLFTGASFAAEPVDPQSGLIWAPGWLNVYQ				
	CNACTSMVTQNRGDKAPFKETLQWMDTQGLWDLSDTWEPTLSYSTYYGRAEMDM				
	TLFRRLPLDASQLPPPAKPAKGSQ"				
	3873..4208				
	/locus_tag="SO0717"				
	3873..4208				


```
/locus_tag="S00717"
/notes="similar to GP.6683510, and SP:P82903; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="monoheme cytochrome c"
/protein_id="AAN53795.1"
/db_xref="GI:24346247"
/translation="MKFIVSLVLCGLSTAVYAEQVQDIERKALATAQCESCHGL
AMIAWAKQFPLNRQCKPAYLNKQLVDFKAGLRDPLMQSQVSALTAEQLQDVALFYSQ
QDIPNLGAN"
4303..4782
/locus_tag="S00718"
/locus_tag="S00718"
4303..4782
/locus_tag="S00718"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN53796.1"
/db_xref="GI:24346248"
/translation="MEWALINNRKIASLIGVLFLPVMGYALQAEVSEANHCQPVSGS
EQHSGQCHNANNAYDKFDGQVYVRGINAKPVTLEIIGQDRYRMTVDGLVTLGELN
TERFQEDENASLFLNMQQPEAEQIKFVKLSKSHSVLRVDANGQLDNTATLQNIQ"
complement(4923..7166)
/locus_tag="S00719"
complement(4923..7166)
/locus_tag="S00719"
/notes="identified by match to PFAM protein family HMM
PF00593"
/codon_start=1
/transl_table=11
/product="iron-dependent receptor, putative"
/protein_id="AAN53797.1"
/db_xref="GI:24346249"
/translation="MKPLPLSLCALAVASALPVMAERQAKPLINEVIVVSGSRMEQNI
EQVGSILVIDEABIERMNSNDFGLSPNESAVDIKGGAGKPTSVTIRIGGNFVMVY
KQGVNVNVOVASPLGPNAGETGRGLTEVQSLKQIEVVKAAASTMYGSDALGGVVMRT
KDAADYLRGSDYVSLNAGSYGNNESYSGGTSAAYSDFENLLTYOHRKGEELPNYP
DTQESDLLQDLSLFKSKRYNHEHTVQLTDLYLEQTLERWEDKLNRAQVPERIDTD
RQTVINGALQSLDKATYLDHDLTVAYFGQTEQLEHRYDFDLRTAARPLTEYRDYE
FKDQRLGLTSFYKIGAEYGHOLTYGMDYQWESQMMRRNNIASNNGDWQTEPTPT
FADTDSNRLGLFDQDITLLNGDILNGLLRDYFRNSPDQQAWEAKGADPADFVMS
DTWSPKVLGLVLTDTVSAYAYGYKMPDQKNGELEVDGSGMTDIIQANYDL
KSEQSHTEWELGIRNADHTQYELTDFDTHAVDPSDQVYFSPKPPVFPVQMPFEFKYQ
YFNKDQVTLGAQAQAHWLTDSIELGNISYTHGTDEGNQNLNSVPLKNGLVNMY
SDIGMETDTCGVVRWADKMDRTDNLITDDIICVMGQCIPKESFNEVYNTAGYAVLD
LTMLGLVTDNLHLRAGVYNLPDKYVDYDVAGQSVFLNLSNWNLQKDDFTQPGRYFN
LSMNYQF"
complement(7394..7903)
/locus_tag="S00720"
complement(7394..7903)
/locus_tag="S00720"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN53798.1"
/db_xref="GI:24346250"
/translation="MRPLTPLLILLPLLVGYQQQAKANELPSKDKLAPQSAOOTIT
PTQAVSPAAKAPADLAQPOLHRVTDKGLTKPNMPTKGEAMVLOCTVYVHLGEGF
WGIADVNGQHLPLKPLPQYRKDGRISLFSQAQETIGMTITQQWGLKSNLSDITVIGQV
ESQSDAPRL"
complement(8096..8347)
/locus_tag="S00721"
complement(8096..8347)
/locus_tag="S00721"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN53799.1"
/db_xref="GI:24346251"
```

```
/translation="MNPILAILKEHNVYDAKIQLFOALTDNPLMAMGIQLGIPPE
KLQQLMALVMQNPALIKAEVLELGLDFAKVEAAKAQLQK"
Query Match 25.7%; Score 382; DB 1; Length 11421;
Best Local Similarity 66.6%; Pred. No. 6.9e-102;
Matches 579; Conservative 0; Mismatches 285; Indels 6; Gaps 2;
QY 623 TCAAAATTTCTACTATCCAGAAACACTGGATTTAACTCTCTCTGAGATTACACAGCCCTGA 682
DB 9113 TAATCAATTTTGGTGGCTGAAATAATTAACCTTAAGCCCATACGCGAAATGCGATTGA 9172
QY 683 ATCAAAATCCCTGGGGGCTGATTTTGTATATGACCACAGATTTCACACAGCTGGGATATGA 742
DB 9173 ATCCAAATCCCTATGGCAGTGAATATCGCTATGCCGAAGCCCTTTAACTTTAGATCTCGA 9232
QY 743 GGCTCTCAAAAGAGATATCAAAAGATTTGCTGACAACTTCCACAGATTGGTCCCTCGCGA 802
DB 9233 CGCGGTTAAAAAGACATTAAGAGCTTTGATGACCGAATCCCAAGATTGGTGGCCGCGGA 9292
QY 803 TTATGGTCATATGGTCTCTTTCTTTATTCGTATGGCTTGGCACGCTGGCCGCGCAATACAG 862
DB 9293 TTATGGTCATATGGACCTTTCTTTATCCGTATGGCTTGGCATAGCGCGGCGGTATACCG 9352
QY 863 GACATATGATGGCGGCGGAGCGCCAGCTGGTGGTCAGCAACGTTTTTGAACCGCTGAACAG 922
DB 9353 GATTTTGTATGGCCGCGGTGCTGCTGTGGTGGTCAACAGCTTTGAGCCTTTAAATAG 9412
QY 923 CTGSC---CGGATAACCTTTAATCTGGATAAAGCCGCTCGATTTGCTGTGGCCAGTCAAGAA 979
DB 9413 CTGSCCAGCAGATAAAGCTGAGCTTGGATAAGGCGCGCGCTTTGTTGGCTATCAAGCA 9472
QY 980 AAAATACGGTCCAGTATTTCTTGGGAGACCTGATGGTCTGTGCTGGTAATGTGGCCT 1039
DB 9473 AAAATACGGCAGCAAGCTTTCTTGGGCGCAGCTTAATGGTGTAAACGGGTACGCTGGCAT 9532
QY 1040 TGAATCCATCGGATTTAAACCGCTGGGATTTTGTGGCGGAGAGAGAGATGATCGGAGTC 1099
DB 9533 AGAGTCGATGGCTTTAAACCTTTGTTGGCGCGGTGCTGTGACAGATTGGGAAGC 9592
QY 1100 GGACCTGGTATCTGGGGGCTCGACAAACAGCCCTCTTTCAGATTAACCGGGATAA---AAA 1156
DB 9593 CGAAATGGTCAATTTGGGGCTCAGAAAAAGCTGCTCGATAACAAAAGACATAATAGAAA 9652
QY 1157 CGGAAATCTTCAGAAACCTCTTGGCGCACGAGATGGGACTTTATTTATGTCAATCTCGA 1216
DB 9653 GGGTGGATTAGCTAAACCCATGGCGCGACCCAAATGGGCTTTAATTTACGTTAACCCAGA 9712
QY 1217 AGGCGCGGTGGAACACAGATCTCTGGCTTCGCGGAAAGATATCAGGGAAGCTTTTTC 1276
DB 9713 AGGCGCGAAGCGGTGCCGATCTTTAGCTCGGCAAGAAATCCCGCATACCTTTGG 9772
QY 1277 ACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATACATTTGG 1336
DB 9773 CCGTATGGCGATGAATGATGAAGAAACCGTTGCTTAATCGCGGTGCTCACACCTTTGG 9832
QY 1337 TAAAGCAGATGGTCAGCGCTCTCTGAAAAATGATTTGGCGCGAGGGCTGATGTGTGACCC 1396
DB 9833 TAAAGCGCACGGCGCACATGATCCCGCAAAATGCGTGGGCGGCGATCGCGCGCGCAGG 9892
QY 1397 TGTGGAGGACGAGGACTGGGATGGAATAATAATGTTGTACAGAAACCGCAATATAC 1456
DB 9893 CATTGAAGACAGAGTCTAGCTTGGAAAAACAAGATGTGGTAAAGGCCCATTCAGAAATAC 9952
QY 1457 CATCACCAGTGGCTTGAAGAGCGCTGGTC 1486
DB 9953 TGTACCAGTGGCTTGGAGGAGCTGGTC 9982
```

```
RESULT 19
CR628337 02/c
WPCOMMENT
Sequence split into 34 fragments LOCUS CR628337 Accession CR628337
Fragment Name Begin End
```

CR628337_00	1	110000	
CR628337_01	100001	210000	
CR628337_02	200001	310000	
CR628337_03	300001	410000	
CR628337_04	400001	510000	
CR628337_05	500001	610000	
CR628337_06	600001	710000	
CR628337_07	700001	810000	
CR628337_08	800001	910000	
CR628337_09	900001	1010000	
CR628337_10	1000001	1110000	
CR628337_11	1100001	1210000	
CR628337_12	1200001	1310000	
CR628337_13	1300001	1410000	
CR628337_14	1400001	1510000	
CR628337_15	1500001	1610000	
CR628337_16	1600001	1710000	
CR628337_17	1700001	1810000	
CR628337_18	1800001	1910000	
CR628337_19	1900001	2010000	
CR628337_20	2000001	2110000	
CR628337_21	2100001	2210000	
CR628337_22	2200001	2310000	
CR628337_23	2300001	2410000	
CR628337_24	2400001	2510000	
CR628337_25	2500001	2610000	
CR628337_26	2600001	2710000	
CR628337_27	2700001	2810000	
CR628337_28	2800001	2910000	
CR628337_29	2900001	3010000	
CR628337_30	3000001	3110000	
CR628337_31	3100001	3210000	
CR628337_32	3200001	3310000	
CR628337_33	3300001	3345687	
Continuation (3 of 34) of CR628337 from base 200001 (CR628337 Legionella pneumophila Len			
Query Match 25.5%; Score 380; DB 1; Length 110000;			
Best Local Similarity 65.4%; Pred. No. 4.7e-101;			
Matches 589; Conservative 0; Mismatches 305; Indels 6; Gaps 2;			
Qy	590	TTCTACCGCTGTAGCCGTGATAAAAAAGAGACTCAAAAATTTCTACTATCCAGAAACACT	649
Db	79632	TTATGCCCATCGCAAGAGACAGCCGAAACCAATCAATACCTAGTGGCTGAAATGCT	79573
Qy	650	GGATTAACTCTCTGAGATTACACAGCCCTGGAATCAAAATCCCTGGGGGCTGATTGA	709
Db	79572	CGATCTTAGCCGCTGGCGCAACCAATGCCACTTCGAATCCTATGGGTGAAAAATTCAA	79513
Qy	710	TTATGCCACCAAGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAAAGATT	769
Db	79512	CTATGCCGAAGATTTCAATAGCTTGATTTGAATGCCGTGATAGAGATCTCAAAAATTT	79453
Qy	770	GCTGACAACTTCCAGGATTTGGTCCCTGCGGATTTATGGTCATTTATGGTCTCTTTAT	829
Db	79452	AATGACTACCTCGCAAGACTGGTGGCTGCTGATTTATGTTAATGTTCCATTTATCAT	79393
Qy	830	TCGATATGGCTGGCAGCGGTGCGGAAATACAGACATATGATGCCGGGAGGCGCAG	889
Db	79392	ACGATGTGATGGCATGGCGGAGAACTTACCGCATCTATGTCGTGCTGGTGAGCAAA	79333
Qy	890	TGGTGGTCAGCAACGTTTTTGAACCGTGAACAGCTGGCGGATTAACGTTATCTGGATAA	949
Db	79332	TGGTGGCTTCAGCGTTTCGCCCCGCAAAAACAGCTGGCGGATTAATGCCAATCTGGACAA	79273
Qy	950	AGCCCGTCGATTTGTCGGCAGTCAAGAAAAATACCGCTCCAGATATTTCTCTGGGAGA	1009
Db	79272	AGCCAGAGTTTACTGTGGCCAAATTAACGAGATATGCGCGCAAAAATTTTCATGGCTGA	79213
Qy	1010	CCTGATGGTCTGACTGGTAAATGTTGGCCCTTGAATCCATGGGATTTAAACGCTGGGATT	1069
Db	79212	TTTGTGTGTGCTGTGCGAAATGTTGCGATGGAATCTATGGGCTTCAAGACCAATGGGATT	79153
Qy	1070	TGCTGGCGGAGAGAGATGATCTGGAGTCCGAGCTCGATATCTATCTGGGGGCTCGAACAA	1129

Db	79152	TGCTGGAGGCGGTGAAGACGCATGGGAGGCCATCAATATCAATTGGGGCACTGAGGCCAA	79093
Qy	1130	GCCTCTTGAGATAA---CGGCGATAAACCGGAAACTTCAGAACTCTTTCGCGGCAC	1186
Db	79092	ATGGCTGGAAGTAAGCGCCAGATAAAGTTGGAATACTCGAAAAACCGCTTGTCTGGAC	79033
Qy	1187	GCAGATGGGACTTATTTATGTCAATCTGAAGCCCGGTGGAAAAACAGATCCTCTGGC	1246
Db	79032	CGTAATGGGCTTAATCTATGTGAATCGGAAGGACCAACCGCGTTCACAGATCCTCTTGC	78973
Qy	1247	TTCCGCGAAAGATATCAGGAACTTTTTCACGTATGGCCATGATGATGATGAGAGACTGT	1306
Db	78972	TGCGGCGGAAAAAATTCGCGAGACTTTTCGGGCGGTATGGCCATGAATGACGAAGAAACAGT	78913
Qy	1307	GGCCCTGATCGCGGAGGCGATACATTTGTTAAAGCACATGGTGCAGCGCTCTCTGAAA	1366
Db	78912	TGCTTTAAATTTGGCGGCGGCGACGCAATTCGGAANAACACATGGAGCCCGGT---CCGCAA	78856
Qy	1367	ATGTATTGGCGCAGGGCTGATGTCACCTGTGGAGGAGCAGGACTGGGATGGAAAAA	1426
Db	78855	ATATCTGGGCGCAGCGCCAGAGACAGCGGCGCATAGAAGAACAGGCTTTGTTGGAAAAA	78796
Qy	1427	TAAATGTGTACAGGAAACCGCAAAATATACATCACACAGTGGCTTGAAGAGCTCTGTC	1486
Db	78795	CAGTTAGCGCAGCGCAAGGAAAGATACGATTACCAGTGGTTTAGAAGGTGCTCGAC	78736
RESULT 20			
ECIS91TP/c 1829 bp DNA linear BCT 07-JUL-2002			
LOCUS Escherichia coli plasmid pSU2911 including tnpA gene for			
DEFINITION transposase and ORF121.			
ACCESSION X17114.5 S79575			
VERSION X17114.5 GI:21038833			
KEYWORDS insertion sequence; tnpA gene; transposase.			
SOURCE Escherichia coli			
ORGANISM Escherichia coli			
REFERENCE 1 (bases 1 to 1829)			
AUTHORS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
TITLE Escherichia coli			
JOURNAL Enterobacteriaceae; Escherichia.			
MEDLINE 92138630			
PUBMED 1310503			
REFERENCE 2 (bases 1 to 1829)			
AUTHORS Mendiola,M.V. and de la Cruz,F.			
TITLE 1891 transposase is related to the rolling-circle-type replication			
JOURNAL proteins of the pUB110 family of plasmids			
MEDLINE 92335017			
PUBMED 1321417			
REFERENCE 3			
AUTHORS de la Cruz,F.			
TITLE Direct Submission			
JOURNAL Submitted (06-NOV-1989) F. de la Cruz, Departamento Biologia			
REMARK Molecular. Un., Cantabria, Facultad Medicina, C/Cardenal Herrera			
REFERENCE 4			
AUTHORS de la Cruz,F.			
TITLE Direct Submission			
JOURNAL Submitted (27-OCT-2000) F. de la Cruz, Departamento Biologia			
REMARK Molecular. Un., Cantabria, Facultad Medicina, C/Cardenal Herrera			
REFERENCE 5 (bases 1 to 1829)			
AUTHORS Garciaillan-Barcia,M.P.			
TITLE Direct Submission			
JOURNAL Submitted (23-APR-2002) Garcillan-Barcia M.P, Biologia Molecular,			
niversidad de Cantabria, C/ Herrera Oria s/n, Santander, Cantabria,			
39011 SPAIN			

```
COMMENT      On May 20, 2002 this sequence version replaced gi:20376803.
FEATURES
  source
    1..1829
      /organism="Escherichia coli"
      /mol_type="genomic DNA"
      /db_xref="taxon:562"
      /plasmid="psu2911"
  repeat_region
    1..1829
      /insertion_seq="IS91"
  -35_signal
    27..32
      /notes="ORF 121"
  -35_signal
    38..43
      /notes="ORF 121"
  -10_signal
    50..55
      /notes="ORF 121"
  -10_signal
    60..65
      /notes="ORF 121"
  RBS
    79..82
      /notes="ORF 121"
  CDS
    89..454
      /notes="ORF 121"
      /transl_table=1
      /product="hypoetical protein"
      /protein_id="CAC14584.1"
      /db_xref="GI:11071485"
      /db_xref="GOA:Q9F440"
      /translation="MARSAPRRKPKAPORSKLPYVVKLVHPDDFDEDAEVLRFDN
FYDAVECCADLNIPFVFDAGNKKLVFVFRVDDGYPEIARCTEREFEATILAGISAGG
MYCPEGCTVHPDGVPPPP"
    303..1641
      /genes="tnpA"
    303..308
      /genes="tnpA"
    326..331
      /genes="tnpA"
    372..378
      /genes="tnpA"
    396..401
      /genes="tnpA"
    454..1641
      /genes="tnpA"
      /codon_start=1
      /transl_table=11
      /product="transposase"
      /protein_id="CAD30457.1"
      /db_xref="GI:21038834"
      /db_xref="GOA:Q8KMX5"
      /db_xref="UniProt/TrEMBL:Q8KMX5"
      /translation="MLPRADIQQGNRLNWLLEKQPEGSVRPVVTESVTKIMACGTT
LMGYTQCCSPDCCHTKVKYCFKRSRCPHCGVKAGQWIOYLLSLVDPDPMQHIYF
TLPCQYSLVFNHRLLAEMSRIAADVILEICHQDVEPGIFTVIHTWRDQDQWPHI
HLSHTAGVTSHTGNLKLHFKYARKVSMWRYRI TRLLSRKY PELVIPDELA VEGNSKR
DNKCFLDTHVRCGMWNI SRVMDNATHVAVFGSYLKKPVPMSRLHLYAGQDEIGLR
YNSHRTKREYLLMSGDEPNERSHWADKGFWRVRYGFLSPVKRLLLEEVVIVITE
TVKTAQWIRWRGMYORLLKLVDPKLCILGSGQMRFTGLKRGYRLAELVLMHLARQQ
VCG"
  ORIGIN
    Query Match      24.7%; Score 368; DB 1; Length 1829;
    Best Local Similarity 96.0%; Pred. No. 6.7e-99;
    Matches 388; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
  QY      4 CAGTCCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCCATG 62
  DB      626 CAGTCCGGTACCAACAGACCAATGGGTGTAGCCCATCAGCGTGGTCCCGATGCCATG 567
  QY      63 ATTTTGTCA CAGACTCAATGACTACCGGACGCACTGAACTTCCCGTGTGTTTCTCCAGC 122
  DB      566 ATTTTGTCACTGACTCAGTCAACACCGGACGCACTGAACTTCCCGCTGTTTCTCCAGC 507
  QY      123 CAGTTAAGCAGCGGGTTTCCTCTGTAATAATGTCCGCAAAACGGGGAGCATCAGNAGG 182
```

```
|||||
506 CAGTTAAGCAGCGGGTTACCTCTGTAATAATGTCCGCAAAACGGGGAGCATCAGNAGG 447
|||||
183 GCGGGGAAACTCCGTCCGGCCAGTGAACCGTGCACACTCCGGGCACTACATCCGCCGG 242
|||||
446 GCGGGGGGACTCGCTCCGGCCAGTGAACCGTGCACACTCCGGGCACTACATCCGCCGG 387
|||||
243 CGCTGATACCGGCAAGAAATGGTCGCAAACTCCCGCTCCGTCGACGCGGCTATTTTCAGGAT 302
|||||
386 CGCTGATACCGGCAAGAAATGGTCGCAAACTCCCGCTCCGTCGACGCGGCTATTTTCAGGAT 327
|||||
303 ACCCTTCGTTCATCAACACAGTACAAACAGAGAACAGTCTTTTGTCTGACATCCACAA 362
|||||
326 ACCCTTCGTTCATCAACACAGTACAAACAGAGAACAGTCTTTTGTCTGCGCATCCCAA 267
|||||
363 AGAAGGGAATATTTACGTTCTGCGCAGCACTCAACGGCATCGTCA 406
|||||
266 AGAAGGGAATATTTACGTTCTGCGCAGCACTCAACGGCATCGTCA 223
|||||
RESULT 21
AR098264
LOCUS      2238 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION      Sequence 7 from patent US 6074860.
ACCESSION      AR098264
VERSION        AR098264.1 GI:12807521
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unclassified.
REFERENCE      1 (bases 1 to 2238)
AUTHORS      Robertson,D.E., Sanyal,I. and Adhikary,R.S.
TITLE        Catalogues
JOURNAL      Patent: US 6074860-A 7 13-JUN-2000;
FEATURES
  source
    1..2238
      /organism="unknown"
      /mol_type="unassigned DNA"
  ORIGIN
    Query Match      24.0%; Score 357.6; DB 6; Length 2238;
    Best Local Similarity 66.7%; Pred. No. 9e-95;
    Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;
  QY      683 ATCAAAATCCCTGGGGGGCTGATTTTGAATTATGTCACACAGATTTCAACAGCTGATATGGA 742
  DB      168 ATCGGACCCCAACGACCCCGGATTTTGACTATGCGGAAGAGTTTAAAGAAGCTAGATCTGGC 227
  QY      743 GGCTCTGAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTCGGA 802
  DB      228 AGCGGTTAAAAAGGACCTGGCAGCGCTAATGACAGATTCACAGGACTGGTGGCCAGCAGA 287
  QY      803 TTATGTCATTATGTCCTTTCTTTTATGTCATGGCTTGGCAGCTGCGGAAACATACAG 862
  DB      288 TTACGTCATTATGGCCCTTCTTTATGTCATGGCTGGCAGCGCCGACCTACCG 347
  QY      863 GACATATATGCGCGGAGGCGCAGTGGTGGTCAACACGTTTGAACCGCTGAACAG 922
  DB      348 TATCGGTGATGGCGGTGGTGGCTCCGGCTCACAGCGCTTCGCGCTCTCAATAG 407
  QY      923 CTGGCCGGATACGTTAATCTGGATAAAGCCCGTCGATTCGTGGCCAGTCAAGAAAAA 982
  DB      408 CTGGCCAGACAATTCGCAATCTGGATAAAGCAGCTTGTCTTTGGCCCACTCAACAAAA 467
  QY      983 ATACGCTCCAGTATTTCTTGGGGAGACCTGATGGTCCCTGACTGGTAATGTTCCCTTGA 1042
  DB      468 ATACGTCGAATAATCTCTTGGCGGATCTAATGATCTACAGAGAAACGTAGCTTGA 527
  QY      1043 ATCCATGGGATTTAAACCGCTGGGATTTGCTGGCGAAGAGAGATGACTGGGAGTC--- 1099
  DB      528 AACTATGGGCTTTAAACCTTTTGGTTTTGTCAGGTGGCAGAGCAGATGATATGGGAGCTGA 587
  QY      1100 GGNACCTGGTATACTGGGGGCGCTGACAAACAGCCTCTTGCAGATAACGGGATAAA---AA 1156
```

Db	588	AGAGATGTATCTGGGGAGCAGAAACCGAATGGCTGGGAGACAAGCGCTATGAAGGTGA	647
Qy	1157	CGGGAACCTTCAGAAACCTTCTGCCGACGAGATGGGACTTATTATTGTCAATCCTGA	1216
Db	648	CCGAGAGCTCGAAATCCCTGGGAGCCGTACAAATGGGACTCATCTATGTAAACCCCGA	707
Qy	1217	AGCCCGGTGGAAACCAAGATCCTCTGGCTTCGCGGAAAGATATCAGGGAGCTTTTTC	1276
Db	708	AGACCCAAACGGAACCGACACCTATCGCTGCTGCGGTGATATTCGTGAGACTTTTGG	767
Qy	1277	ACGTATGGCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATACATTTGG	1336
Db	768	CGAATGGCAATGAATGACGAGAAACCGTGGCTCTCATAGCGGTGGACACACCTTCGG	827
Qy	1337	TAAAGCACATGGTGACGGCTCTCTCGAAATGATTTATGGGCGAGGCCCTGATGGTCACC	1396
Db	828	AAAAACCATGTGTCTGCCGATCGGAGAAATATGTGGCCGAGAGCCTGCCGCCGAGG	887
Qy	1397	TGTGGAGGACGAGGACTGGATGGAAATAAATGTTGTACAGGAAACGGCAAAATATAC	1456
Db	888	TATTGAAGAAATGAGCTGTGGGTGGAAACACCTACGGCACCGGACAGCGTCCGGATAC	947
Qy	1457	CATCACCAGTGGCTTGGAGGAGCCCTGGTCGA	1488
Db	948	CATCACCAGTGGACTAGAGGCCCTGGACCA	979
RESULT 22			
AR215252			
LOCUS			
Sequence 7 from patent US 6410290.			
AR215252			
VERSION			
AR215252.1 GI:23313419			
KEYWORDS			
Unknown.			
ORGANISM			
Unclassified.			
REFERENCE			
1 (bases 1 to 2238)			
Robertson,D.E., Sanyal,I. and Adhikary,R.S.			
TITLE			
Catalases			
JOURNAL			
Patent: US 6410290-A 7 25-JUN-2002;			
FEATURES			
source			
1..2238			
/organism="unknown"			
/mol_type="mRNA"			
ORIGIN			
Query Match 24.0%; Score 357.6; DB 6; Length 2238;			
Best Local Similarity 66.7%; Pred. No. 9e-95;			
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;			
Qy	683	ATCAAAATCCCTGGGGGCTGATTTTGATATGCCCACAGATTTCAACAGCTGGATATGGA	742
Db	168	ATCGGACCCAAACGCCGATTTTGACTATGCGGAAGATTTTAAAGAAGCTAGATCTGGC	227
Qy	743	GGCTCTGAAAAGATATCAAGATTTCGTGACACTTCCACAGATTGGTCCCTGGGA	802
Db	228	AGCGGTAAAGAGACCTGGCAGCGCTAATGACAGATTACAGGACTGGTGGCCAGAGA	287
Qy	803	TTATGGTCATTATGGTCTTTCTTTATTCGTATGGCTTGGCAGGCTGCCGGAACATACAG	862
Db	288	TTACGGTCATTATGGCCCTCTTTATACGATGGGTGGCAGCGCCGACCTACCG	347
Qy	863	GACATATGATGGCCGGGAGCGCCAGCTGTGTGTCAGCAACGTTTGAACCGCTGAACAG	922
Db	348	TATCGGTGATGGCGGTGGTGGCGTGGCTCCGGCTCACAGCGCTTCGCGCTCTCAATAG	407
Qy	923	CTGGCCGGATAAGCTTAATCTGATAAGCCCGTCGATTCCTGTGGCCACTCAAGAAA	982
Db	408	CTGGCCAGACAATGCCAATCTGGATAAGACCGCTTGCTCTTTGGCCCAACAAACAA	467
Qy	983	ATACGGCTCCAGTATTTCCTGGGGAGACCTGATGGTCTCTGACTGGTAAATGTTGCCCTTGA	1042

Db	468	ATACGCTCGAAAAATCTCTCGGCGGATCTAATATGATACTCACAGGAAACGTAGCTCTGGA	527
Qy	1043	ATCCATGGATTTAAACCGCTGGGATTTCTCGCGGAAGAGAGATGACTGGGAGTC---	1099
Db	528	AACTATGGGCTTTAAACACTTTTGGTTTTGCAGGTGGCAGCAGATGTATGGGAGCCTGA	587
Qy	1100	GGACCTGGTATATCTGGGGCCCTGACAAACAGCCTCTTTCAGATAAACCGGATAAA---AA	1156
Db	588	AGAAGATGTATCTGGGAGCAGAAACCGAATGGCTGGGAGACAAGCGCTATCAAGGTGA	647
Qy	1157	CGGGAACCTTCAGAAACCTTTTCCGCCACGAGATGGGACTTATTATGTCAATCCTGA	1216
Db	648	CCGAGAGCTCGAAATCCCTCGGAGCCGTACAAATGGGACTCATCTATGTAAACCCCGA	707
Qy	1217	AGCCCGGTGGAAACCAAGATCCTCTGGCTTCGCGAAGATATCAGGGAGCTTTTTC	1276
Db	708	AGGACCCAAACGGAACGACAGCCCTATCGCTGCTGCGGTGATATTCGTGAGACTTTTGG	767
Qy	1277	ACGTATGGCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATACATTTGG	1336
Db	768	CGAATGGCAATGAATGACGAGAAACCGTGGCTCTCATAGCGGTGGACACACCTTCGG	827
Qy	1337	TAAAGCACATGGTGACGGCTCTCTGAAATAATGTTATGGCGCAGGCCCTGATGGTCACC	1396
Db	828	AAAAACCATGTGTCTGCCGATCGGAGAAATATGTGGCCGAGAGCCTGCCGCCGAGG	887
Qy	1397	TGTGGAGGACGAGGACTGGGATGGAAATAAATGTTGTACAGGAAACGGCAAAATATAC	1456
Db	888	TATTGAAGAAATGAGCTGTGGGTGGAAACACCTACGGCACCGGACACGGTCCGGATAC	947
Qy	1457	CATCACCAGTGGCTTGGAGGAGCCCTGGTCGA	1488
Db	948	CATCACCAGTGGACTAGAGGCCCTGGACCA	979
RESULT 23			
AE016983			
LOCUS			
Shigella flexneri 2a str. 2457T section 6 of 16 of the complete			
genome.			
AE016983 AE014073			
VERSION			
AE016983.1 GI:30041148			
KEYWORDS			
Shigella flexneri 2a str. 2457T			
SOURCE			
Shigella flexneri 2a str. 2457T			
ORGANISM			
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
Enterobacteriaceae; Shigella.			
REFERENCE			
1 (bases 1 to 290628)			
Wei,J., Goldberg,M.B., Burland,V., Venkatesan,M.M., Deng,W.,			
Fournier,G., Mayhew,G.F., Plunkett,G. III, Rose,D.J., Darling,A.,			
Mau,B., Perna,N.T., Payne,S.M., Runyen-Janecky,L.J., Zhou,S.,			
Schwartz,D.C. and Blattner,F.R.			
Complete Genome Sequence and Comparative Genomics of Shigella			
flexneri Serotype 2a Strain 2457T			
Infect. Immun. 71 (5), 2775-2786 (2003)			
JOURNAL			
12704152			
PUBMED			
2 (bases 1 to 290628)			
REFERENCE			
Wei,J., Goldberg,M.B., Burland,V., Venkatesan,M.M., Deng,W.,			
Fournier,G., Mayhew,G.F., Plunkett,G. III, Rose,D.J., Darling,A.,			
Mau,B., Perna,N.T., Payne,S.M., Runyen-Janecky,L.J., Zhou,S.,			
Schwartz,D.C. and Blattner,F.R.			
Direct Submission			
Submitted (13-JUN-2002) Genetics Laboratory, University of			
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA			
FEATURES			
source			
1..290628			
/organism="Shigella flexneri 2a str. 2457T"			
/mol_type="genomic DNA"			
/strain="2457T"			
/serotype="2a"			
/db_xref="taxon:198215"			
93..1085			
gene			

```

/feature="teba"
/locus_tag="S1493"
93..1085
/feature="teba"
/locus_tag="S1493"
/function="transport; Drug/analog sensitivity"
/notes="residues 1 to 330 of 330 are 94.24 pct identical to
residues 1 to 330 of 330 from Escherichia coli K-12 ;
B1429"
/codon_start=1
/transl_table=11
/product="Tellurite resistance protein TehA"
/protein_id="AAP16878.1"
/db_xref="GI:30041149"
/translation="MQSDKVLNLPAGYFGLVGTIGMGPWRYASQWVQSHWLGDL
VLAMIWGLTSAFIRLIRFPHSPHVAEVRHPHVSFLFPATTMLVAIGFVFWFR
PLVCMFSPGVVGLAVAAWQTAGLWRGSHPEATTGLYLPVANNFISAMACGALG
YTAGLVLAGVFSWLSLEPVLIRLSRGELPTALRISIGIQALPACVACSAWLSV
NGRGDPTLAKMLFCYGLQLLFLMLRMPWYLSQPNFASFWFSFGVSALATTGLHGS
GSDNGFFHTLAVPLFIPTNFIALLIRTFALLMQGKLLVYTERAVLMKAEDKE"
1082..1675
/feature="tebB"
/locus_tag="S1494"
1082..1675
/feature="teba"
/locus_tag="S1494"
/function="putative transport; Drug/analog sensitivity"
/notes="residues 1 to 197 of 197 are 98.98 pct identical to
residues 1 to 197 of 197 from Escherichia coli K-12 ;
B1430"
/codon_start=1
/transl_table=11
/product="Tellurite resistance protein TehB"
/protein_id="AAP16879.1"
/db_xref="GI:30041150"
/translation="MIRDENYFDKYELRTHSEVLEAVKVKPGKTLDLGCGNGRN
SLYLAANGYDVDAKMAKIANVERIKSIENLNLNLTTRVVDLNLNLPDGOYDPIST
VLMLFLAKTIPGLIANMQRTKPGYNLIVANDTADYPTCTGTFAPFEGELARRY
EGMRYKNEVDGELHRTDANGRIKURFATMLARKK"
1978..2634
/locus_tag="S1495"
1978..2634
/locus_tag="S1495"
/notes="residues 1 to 218 of 218 are 91.44 pct identical to
residues 1 to 222 of 222 from Escherichia coli K-12 ;
B1431"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP16880.1"
/db_xref="GI:30041151"
/translation="MRTTSFAKVAALCGLLALSCASKITQPKYNNYSDLKETTSAT
GKPLRVWDPSPDQSKYDSIVNPIYYPVPKSTQGVKVLKILNLTWNTMEKEAIA
QRKPVVTTAGPRSLIFRGATGVDTSGEGLQFVBPVVALVAVCTOMATGHRTWDTRL
YFEGELIDAATNKPKVIKVRQGECKDLNNESTPNAFENIKQVIDDMATDMPDVNKK"
3226..3486
/locus_tag="S1496"
/notes="point mutation"
/pseudo
4059..5886
/feature="insertion element"
/insertion_seq="IS91"
complement(4247..5434)
/locus_tag="S1497"
complement(4247..5434)
/locus_tag="S1497"
/function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 388 of 395 are 97.93 pct identical to
residues 1 to 388 of 426 from GenPept : >emb|CAA34970.1|
(X17114) transposase IS91 [Escherichia coli]; partial"
/codon_start=1
/transl_table=11

```

```

/product="IS91 orf"
/protein_id="AAP16881.1"
/db_xref="GI:30041152"
/translation="MLPRFADIFQQGNRWLNLWLEKQPEGSVRPVVTVSVTKIMACGTT
LMGYTQCCSSPDCCHTKVCFRCKSRSCPHGCVKAGQWQIYLLSLVDPDQWQHIVP
TLPCQVSLVFNHRWLLAEMRIAADVILSICHQTDVEPGITVITHWGRDQWHPHI
HLSTAGVTSVSGHTKNLHFYARKVMSMRYRITRLLSRKYPELAVIPELAVEGNSKR
DWRPLDTHYRGWNVNVRMDNATHVAYFGSYLKKPVPMSRLSHYAGQDDIGLR
YNSHRTKREYLLMGDFMERFESHWADGFRWRYIGFLSPAKRLLLEVVYITE
TVAKTAQITWRGMYQLLKVDPLKCVLQSKQMRFTGLKRGYRLAEQVLMHBP LARMR
WCG"
5463..5789
/locus_tag="S1498"
5463..5789
/locus_tag="S1498"
/function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 67 of 108 are 98.50 pct identical to
residues 1 to 67 of 79 from GenPept : >gb|AAC70084.1|
(AF074613) unknown [Escherichia coli O157:H7]; partial"
/codon_start=1
/transl_table=11
/product="IS91 orf"
/protein_id="AAP16882.1"
/db_xref="GI:30041153"
/translation="MNRATLRVHAAGADTGKNGRKLPLRAAGYFRIPFVINTVKPED
QLFVSGIHKEGNIQVCALNGIVIKIQAQNFVFFVKIIRMKPHNIAGKPEASLWCGF
AFTGFS"
6302..7069
/feature="insertion element"
/insertion_seq="IS1X1"
6357..6632
/locus_tag="S1499"
6357..6632
/locus_tag="S1499"
/function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 91 of 91 are 100.0 pct identical to
residues 1 to 91 of 91 from GenPept :
>gb|AAK18550.1|AF348706_239 (AF348706) iso-IS1 orfa
[Shigella flexneri]"
/codon_start=1
/transl_table=11
/product="IS1 orfa"
/protein_id="AAP16883.1"
/db_xref="GI:30041154"
/translation="MASISRCSCSATEGVWNGKSTAGHQRYLCSHCRTKWLQFT
YTASQPTHQKIIDMAMNGVCRASARIMGVGLNTVLRHLKNSGRSR"
6551..7054
/locus_tag="S1500"
6551..7054
/locus_tag="S1500"
/function="IS, phage, Tn; Transposon-related functions"
/notes="residues 1 to 167 of 167 are 89.82 pct identical to
residues 1 to 167 of 167 from GenPept :
>gb|AAK18365.1|AF348706_54 (AF348706) IS1 transposase
[Shigella flexneri]"
/codon_start=1
/transl_table=11
/product="IS1 orfB"
/protein_id="AAP16884.1"
/db_xref="GI:30041155"
/translation="MSCQCTHYGRWPQHGTSLKLRPQSVTSRISQPSDVIVCAEMD
ERWGYCAKSRQRLWFAYDRITRTVVVHVHFGERTLATLRLLSLLSAFVWMTDG
WPLYERSLKGLHVIKRYTQRIERHNLRLQHLARLGRKLSLSPKSVSLVHLDHVGHY
LNIRKHYQ"
7331..7783
/feature="ppba"
/locus_tag="S1501"
/notes="truncated"
/pseudo
7781..9044
/feature="insertion element"
/insertion_seq="IS600"
7845..8147

```


TLAFTLGSVLLTLFLMPPLFRAGKSGQNLTHLRGQYRQQLTAWLQGOAELTIFGA
SDRYRTOLENTIOWLEAQREOSLTAISOAILLIGLALVILWMAWGGVGNAP
GALLALVFCALAFALAPVYGAFLHGVIAVRLSDLDKQPEVTFDTQTRVA
QVSSITLRDVFYTPESQKALGKISLVNAGEHIALGRGCGKSTLLQLLTTRAWDP
QOGEILLNDSFASLNEALRQTTISVVPQRVHLFSATLRDNLASPGSDALSEIL
RVGIELLEDGKGLSGGELRLAIARTLLHDAPLVLDDETEGLDAT
TESQILLEAKMREKTVLMYTHRLGRLSRFQIIVMDNGQIIEQGTHABELARQRY
YORKQGL"
complement (3194..4960)
/gene="cydb"
/locus_tag="S0887"
complement (3194..4960)
/gene="cydb"
/locus_tag="S0887"
/function="transport; Not classified"
/note="residues 1 to 588 are 92.85 pct identical to
residues 1 to 588 from Escherichia coli K-12 :
B0887"
/codon_start=1
/transl_table=1
/product="ATP-binding component of cytochrome-related
transport"
/protein_id="AAP16351.1"
/db_xref="GI:30040620"
/translation="MNSROKELTRWLKQSVISQRMNLSRLGFLVSGILIIAQMWF
MARLOHIMENIPREALLPFTLLVTLFVLRVWMLRERVGHAGQHRRFAIRQV
LDRLQAGPAWIOKPKAGSNATLVEQIDMDHYRIYLPOMALAVSVPLIIIVAIFF
SNMAAILIGTAPLIFLWALVGMGAADANRRNLALARSLGHFLDRLGEMETIRF
GRGAETESRSASEDFRMEVRLFAFLSSGILEFTLSIALVAVYFGFYLIGL
DFHGYGVTLADGFLAILLAPEFPQPLRLDGTYHAKAQAAGVADSLKTFMPTLAH
PORGEALATDPLTIAEDLFTSPGKTLAGPLNTLPAGORAVLVGSGSKSL
NLALSGLSVQSGIRINGIELRDLSPESWRKHLVWQONQPLPAKLRDNLVLRPDA
SEQLQALONAWSEFLPLPQGVDPVGDQARLSVGOQRVAVARALINPLCSLLL
LDEPAASLDHSEQRVMEALNAASLRQTLTMVTHQLEDLADMDVIMVMDGRITTEQR
YAELSVAGGFPATLLAHQEEI"
complement (5083..6048)
/gene="trxB"
/locus_tag="S0888"
complement (5083..6048)
/gene="trxB"
/locus_tag="S0888"
/function="enzyme; 2-Deoxyribonucleotide metabolism"
/note="residues 1 to 321 are 95.95 pct identical to
residues 1 to 321 of 321 from Escherichia coli K-12 :
B0888"
/codon_start=1
/transl_table=1
/product="thioredoxin reductase"
/protein_id="AAP16352.1"
/db_xref="GI:30040621"
/translation="MGITTKSKLLILGSGPAGYTAAYVAAANLQPLVITGMKGGOL
TTTTVENWPGDNDLGTPLLMARMHEHATKFEETIFDHINKVDLQNRFFRLNGDNG
EYTCDALIATGASARYLGFSEAPKRGVVSACATCDGFFYRNQKVAVGNGTAVE
EALYLSNAGEVLIHRRDGFRAEKILIKRLMDKVNGNIIILHNRITLBEVTGDMGV
TGVLRLDTONSDNIESDVLGFLVAIGHSPNTAIFEGQLENGYIKVQSGIHGNATQ
TSIPGVPAAGDVMDHITRQAITSGTCCMAALDARYLDGLADAK"
6592..7086
/gene="lrp"
/locus_tag="S0889"
6592..7086
/gene="lrp"
/locus_tag="S0889"
/function="regulator; Central intermediary metabolism:
Pool, multipurpose conversions"
/note="residues 1 to 164 are 100.00 pct identical
to residues 1 to 164 of 164 from Escherichia coli K-12 :
B0889"
/codon_start=1
/transl_table=1
/product="regulator for leucine (or lrp) regulon and
high-affinity branched-chain amino acid transport system"
/protein_id="AAP16353.1"
/db_xref="GI:30040622"

/translation="MVDSKRPKGLDLRDLRNLILNELOKGRISNLSKRVGLSPPT
CLERVRLERQGFQGTALNLNPHYLDAISLTLFVEITLNRGAPDVFOFNVAOKLEE
IQECHLVSGDFDYLLKTRVPMDSAYRKLGLLETLRPLFVNDTRTYVWVEEYKQSNRLV
IKTR"
7221..11249
/gene="ftsK"
/locus_tag="S0890"
7221..11249
/gene="ftsK"
/locus_tag="S0890"
/function="phenotype; Cell division"
/note="residues 1 to 737 of 1342 are 68.24 pct identical
to residues 1 to 737 of 1329 from Escherichia coli K-12 :
B0890"
/codon_start=1
/transl_table=1
/product="cell division protein"
/protein_id="AAP16354.1"
/db_xref="GI:30040623"
/translation="MSQETEDKEVTUTKLSGRRLLLEALLILVLFVWMLMAALLSF
NPSPDSQSWATHEPIHNLGMPGAMLADTLFFIFGYMAYTIPVIVGGCWFARHQHS
SDYIDYFAVSLRIIGVLAAILITSCGLAIAINADDIWFASGCVIGLSLLTLQPLLHS
SGGTIALLCVAAAGLTLFTGMSWTIAEKLGLWILNLTASNTRDDTWVDDEDE
DDEYEDENHGHQESRRARILRGALARRKLAKEFNPNGROTDAALFSGKRWDDDE
EITYTARGVAADDDVLFSGNRATQPEYDEYDPLNAPITEPVAVAAAATATQSHA
APPEVPTQTPVASVDVPPSQPTVAMQVFPQTEPVIAPEGEYFQQSQYQAPVQ
VNEFLQVFPQQPYAPAAEPAQPYAPAPQPVAGNAQWAEQOSTFAPLSTVQ
TEQYQAPAAEPLYQQPVEQQPVVEPEVVEETKAPPLPYFVEVEKRRARE
QLAAWQPIPEPVKEPEPKSLKAPSAVAVPVEAAAVSLASGVKKATLATGAAA
TVAAPVFLSNGSGPRQVKEIGPOLPRKRIKRVTRRELASVGLKPSQRAEAKA
REARQNYDSQDQYNDDEIDAMQDELARQACTQQRYGEQYQHDVFNVEDADAAA
EAEALQFQACTQQRYSGEGPAGNPFSLDDFFSPMKALLDDGHPFLPTFPIVEPVQ
QPPQVAPQQYQQPQVAPQQYQQPQVAPQQYQQPQVAPQQYQQPQVAPQQYQQPQV
APQYQQPQVAPQQYQQPQVAPQQYQQPQVAPQQYQQPQVAPQQYQQPQVAPQQYQQP
DLTPPPEVEPVDFTALEQMARLVEARLADFRIKADVNVYSGPVITRPELNLAPGV
KAARISNLSRDLARSLSLTVAVRVVEVPGKPYVGLLELPNKKQRTVYLRVLEVDNAKFRD

Query Match 23.6%; Score 352; DB 1; Length 292504;
Best Local Similarity 93.6%; Pred. No. 1.4e-92;
Matches 378; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

Qy	4	CAGTCCGGAGATGAA-AGCACCACTGCTGTACCCCATCAGCGTGTCCCGCAGGCCATG	62
Db	222848	CAGTCCGGTGTATGAACACACCACTGCTGTAGCCCATCAGTGTGGTCCCGCAGGCCATG	222907
Qy	63	ATTTTGTCTCAGACTCAATGACTTACCGGACGCACTGAACTTCGGCTGTGTTCTCCAGC	122
Db	222908	ATTTTGTCTACTGATCAATTAACCGGACGCACTGCTCCCTTCGGTGTGTTCTCCAGC	222967
Qy	123	CAGTTAAGCCAGCGGTTTCCCTGCTGAAAATGTGCGCAAAAACGGGGAAGCATCAGAAGG	182
Db	222968	CAGTTAAGCCAGCGGTTTCCCTGCTGAAAATATCGCAAAAACGGGGAAGCATCAGAAGG	223027
Qy	183	CGCGGGGAACCTCCGTCGGCCAGTGAACCGTGCACACTCCGGCAGTACATCGCCCGG	242
Db	223028	CGCGGGGCACTCCGTCGGCCAGTGAACCGTGCACACTCCGGCAGTACATCAGCCCGG	223087
Qy	243	CGCTGTATACCGGCAAGAAATGCTCGCAAACTCCGCTCCGTCAGCGGGCTATTTTCAGGAT	302
Db	223088	CGCTGTATACCGGCAAGAAATGCTCGCAAACTCCGCTCCGTCAGCGGGCGGATTTTCGGAT	223147
Qy	303	ACCTTTCGTGTCATCAACACGTCACAAAACGAGAAGACCAAGCTTTTGTCTGACATCCCAA	362
Db	223148	ACCCTTCGTGTCATCAACACGTCACAAAACGAGAAGACCAAGCTTTTGTCTGACATCCCAA	223207
Qy	363	AGAGGGGAATATTCAGGTCTCGCAGCACTCAACGGCATCGTCA	406
Db	223208	AGAACGGGAATATTCAGGTCTCGCAGCACTCAACGGCATCGTCA	223251

RESULT 25
AE015130
LOCUS

AE015130 10206 bp DNA linear BCT 18-OCT-2002

GAAPETNPRLPLDGLGGYIAPDNLTITLSVGHSLFDERFGLAPOMPKLOKWT
FPNLSLDAALCHGDLVLIQICANTQDVTNHALRDIKHTPDLSSVRWRKRGFISHAAR
SKGKETPNLKGFGDGTANPDQSQNDKLMQKVWVWVWADQOEPAWTIGGSYQAVRLIQR
VEFWDRPLKEQQITFGDKQGTAPLGMQHEHDPDYASDPDGKVIADLRILANPR
TAESSESLMRDRGYSYSLGVNNSQGLMDGLLFVCYQHDLEKGFILTVQKRLNGALEEY
VKPIGGYFFSLPGLPKVANDYLGRALLQV"

gene

/gene="phoH"

/locus_tag="SF1022"

6307..7230

/locus_tag="phoH"

/locus_tag="SF1022"

/note="Residues 2 to 307 of 307 are 98 pct identical to
residues 49 to 354 of a 354 aa protein from Escherichia
coli O157:H7 ref: NP_309293.1"

/codon_start=1

/transl_table=1

/product="ATP-binding pho regulon component PhoH"

/protein_id="AAN42648.1"

/db_xref="GI:24051309"

/translation="MRSRFSAFYKANYGLQSPDSVRHNRAANGACDEYKOLKVL
SMGQKAVIKARREAKRLRDSRSHKQREESVTSLVQMSGVEAIGWARDSDRTSPI
LARNEAQHLYLQAIETESQLIPATGEAGCGKTWISAACAALAIHKDVRITVTPVLQ
ADEDLGLPGDIAEKFPAPYFRPVYDILLVRLGASFMOCYCLRPEIKGVETAPFAYMRGR
TFENAVILDBAQNVTAAQMKMFLTRIGENVTVIVNGDIITQCDLPRGVCGLSALER
FEEDMVGIVRFKEDCVRSALCQRTLHAYS"

gene

/gene="tra5_55"

/locus_tag="SF1023"

complement(7340..8206)

/gene="tra5_55"

/locus_tag="SF1023"

/note="Residues 1 to 288 of 288 are 100 pct identical to
residues 1 to 288 of a 288 aa protein from Escherichia
coli ref: NP_053084.1"

/codon_start=1

/transl_table=1

/product="transposase"

/protein_id="AAN42649.1"

/db_xref="GI:24051310"

/translation="MKTVFIEKHQAFSIRKMCRLVRVRSWYTWQRRTRISTROQ
PROCHDSVLAAFPTRSKORYCAPRLTDLRAQGYPPNVKTVAAASLRQGLRAKSRKF
SPVSYRAHGLPVSNLEODFYASGPNQWAGDITLRTDEGLYLAVVIDLMSRAVI
GWSNRPRTAQLACDALQMALWRKRPRNVIVHTRDGGQYCSADYQALKRHLNRGSM
SAGCCYDNACVSEFFHSLKVECIHGFHSREIMRATVFNIECDYNRRWRHSWCGG
LSPEQFENQLA"

gene

complement(8203..8511)

/locus_tag="SF1024"

complement(8203..8511)

/locus_tag="SF1024"

/note="Residues 1 to 102 of 102 are 98 pct identical to
residues 1 to 102 of a 102 aa protein from Escherichia
coli K12 ref: NP_414832.1"

CDS

Query Match 23.5%; Score 350.4; DB 1; Length 10206;
Best Local Similarity 93.3%; Pred. No. 1.8e-92;
Matches 377; Conserved 0; Mismatches 26; Indels 1; Gaps 1;

QY 4 CAGTCCGGAGATGAA-AGCACCACTGTGTGTACCCCATCAGCGTGGTCCCGCAGGCCATG 62
DB 2078 CAGTCCGGGTATGAACACACCACTGGGTGTAGCCCATCAGTGTGGTCCCGCAGGCCATG 2137

QY 63 ATTTTGTGCAGACTCAATGACTACCGGACGCACTGAACCTTCCGGTGTGTTCTCCAGC 122
DB 2138 ATTTTGTGCTACTGATTCAATTACCACCGGACGCACTGCCCCCTTCCGGTGTGTTCTCCAGC 2197

QY 123 CAGTTAGCCAGCGGTTTCCCTGCTGAAATGTCCGGCAAAACCGGAGAGCATCAGAAG 182
DB 2198 CAGTTAGCCAGCGGTTTCCCTGCTGAAATATCGGCAAAACCGGAGAGCATCAGAAG 2257

QY 183 GCGGGGGAATCCCGTCCGGCCAGTGAACCGGTGCCACACTCCGGGACGTACATCCGCCGG 242
DB 2258 GCGGGGGAATCCCGTCCGGCCAGTGAACCGGTGCCACACTCCGGGACGTACATCCGCCGG 2317

QY 243 CGCTGATACCGGCAAGAATGGTCCGAAACTCCCGCTCCGTGACGCGGCTATTTCCAGGAT 302
DB 2318 CGCTGATACCGGCAAGAATGGTCCGAAATTTCCCGCTCCGTGACGCGGCTATTTCCAGGAT 2377
QY 303 ACCCTTCTGTCATCAACACGCTACAAACCCAGAGACACGCTTTTCTGTCATCCCAAA 362
DB 2378 ATCCTTCTGTCATCAACACGCTACAAACCCAGAGACACGCTTTTCTGTCATCCCAAA 2437
QY 363 AGAAGGGAATATTTCAGGTCTGCGCAGCACTCAACCGGATCGTCA 406
DB 2438 AGAAGGGAATATTTCAGGTCTGCGCAGCACTCAACCGGATCGTCA 2481

Search completed: March 11, 2005, 06:54:57
Job time : 8354.22 secs

This Page Blank (uspto)

```
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8750
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8750

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      10 ATCGTCAG 17

RESULT 19
US-09-866-108A-8751
; Sequence 8751, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866.108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8751
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8750

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      10 ATCGTCAG 17
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8751
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8751

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      9 ATCGTCAG 16

RESULT 20
US-09-866-108A-8752
; Sequence 8752, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866.108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8752
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8752

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      8 ATCGTCAG 15
```

```
RESULT 21
US-09-866-108A-8753
; Sequence 8753, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8753
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8754
Query Match 100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 7 ATCGTCAG 14

RESULT 22
US-09-866-108A-8754
; Sequence 8754, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8753
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8753
Query Match 100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 7 ATCGTCAG 14

RESULT 23
US-09-866-108A-8755
; Sequence 8755, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8754
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8754
```

```

; PRIOR FILING DATE: 2001-01-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Acomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8755
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8755

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      5 ATCGTCAG 12

RESULT 24
US-09-866-108A-8756
; Sequence 8756, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: ACOMICA-7
; CURRENT FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Acomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8756
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8756

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      4 ATCGTCAG 11
```

```

RESULT 25
US-09-866-108A-8757
; Sequence 8757, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: ACOMICA-7
; CURRENT FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Acomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 8757
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108A-8757

Query Match      100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      3 ATCGTCAG 10

Search completed: March 11, 2005, 13:12:32
Job time : 3.24051 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:06:25 ; Search time 44.8366 Seconds
(without alignments)
8645.645 Million cell updates/sec

Title: US-09-674-277-1_COPY_400_407

Perfect score: 8

Sequence: 1 atcgctcag 8

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	8	100.0	10	6 AR303606	AR303606 Sequence
C 2	8	100.0	12	6 BD231923	BD231923 Modified
C 3	8	100.0	12	6 AX019721	AX019721 Sequence
C 4	8	100.0	12	6 AX080675	AX080675 Sequence
C 5	8	100.0	14	6 AR011717	AR011717 Sequence
C 6	8	100.0	14	6 AR075714	AR075714 Sequence
C 7	8	100.0	14	6 AR126298	AR126298 Sequence
C 8	8	100.0	14	6 BD132864	BD132864 Nucleic a
C 9	8	100.0	14	6 BD205228	BD205228 Nucleotid
C 10	8	100.0	14	6 179345	179345 Sequence 5
C 11	8	100.0	14	6 AX011314	AX011314 Sequence
C 12	8	100.0	15	6 AR301299	AR301299 Sequence
C 13	8	100.0	15	6 AX103974	AX103974 Sequence
C 14	8	100.0	15	6 AX355608	AX355608 Sequence
C 15	8	100.0	15	6 AX547027	AX547027 Sequence
C 16	8	100.0	15	6 BD064833	BD064833 Method fo
C 17	8	100.0	16	6 A72834	A72834 Sequence 14
C 18	8	100.0	16	6 A72978	A72978 Sequence 14
C 19	8	100.0	16	6 BD205229	BD205229 Nucleotid

C 20	8	100.0	15	6 AR329607	AR329607 Sequence
C 21	8	100.0	16	6 AR329608	AR329608 Sequence
C 22	8	100.0	16	6 AX011315	AX011315 Sequence
C 23	8	100.0	16	6 BD093184	BD093184 A gene co
C 24	8	100.0	16	6 BD093185	BD093185 A gene co
C 25	8	100.0	17	6 CQ624010	CQ624010 Sequence
C 26	8	100.0	17	6 CQ624011	CQ624011 Sequence
C 27	8	100.0	17	6 CQ624012	CQ624012 Sequence
C 28	8	100.0	17	6 CQ624013	CQ624013 Sequence
C 29	8	100.0	17	6 CQ624014	CQ624014 Sequence
C 30	8	100.0	17	6 CQ624015	CQ624015 Sequence
C 31	8	100.0	17	6 CQ624016	CQ624016 Sequence
C 32	8	100.0	17	6 CQ624017	CQ624017 Sequence
C 33	8	100.0	17	6 CQ624018	CQ624018 Sequence
C 34	8	100.0	17	6 CQ624019	CQ624019 Sequence
C 35	8	100.0	17	6 AR188444	AR188444 Sequence
C 36	8	100.0	17	6 AR188445	AR188445 Sequence
C 37	8	100.0	17	6 AR324297	AR324297 Sequence
C 38	8	100.0	17	6 AR324298	AR324298 Sequence
C 39	8	100.0	17	6 AR328924	AR328924 Sequence
C 40	8	100.0	17	6 AR465073	AR465073 Sequence
C 41	8	100.0	17	6 AR465074	AR465074 Sequence
C 42	8	100.0	17	6 AR465075	AR465075 Sequence
C 43	8	100.0	17	6 AR465076	AR465076 Sequence
C 44	8	100.0	17	6 AR465077	AR465077 Sequence
C 45	8	100.0	17	6 AR465078	AR465078 Sequence

ALIGNMENTS

RESULT 1
AR303606/c
LOCUS AR303606 10 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 331 from patent US 6544736.
ACCESSION AR303606
VERSION AR303606.1 GI:31692382
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10)
AUTHORS Shimamoto,A., Furuichi,Y., Shibata,Y., Funaki,H., Ohara,E. and Wataniki,M.
TITLE Method for synthesizing cDNA from mRNA sample
JOURNAL Patent: US 6544736-A 331 08-APR-2003;
FEATURES Location/Qualifiers
source 1..10
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2
RESULT 2
BD231923 12 bp DNA linear PAT 17-JUL-2003
LOCUS BD231923
DEFINITION Modified nucleic acid probes and uses thereof.
ACCESSION BD231923
VERSION BD231923.1 GI:33041693
KEYWORDS JP 2002510465-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
1 (bases 1 to 12)
AUTHORS Weston,A., Asseberg,R., Marsh,P., Mock,G.A., Ray,T.D., Wharam,S.D.

and Cardy, D.L.N.
Modified nucleic acid probes and uses thereof
Patent: JP 2002510465-A 1 09-APR-2002;
CYTOCELL LTD
COMMENT OS Artificial Sequence
PN JP 2002510465-A/1
PD 09-APR-2002
PF 26-JAN-1999 JP 2000528713
PR 27-JAN-1998 GB 9801628.0, 29-APR-1998 GB 9809014.5 PI
ANTHONY WESTON, RENE ASSENBERG, PETER MARSH, GRAHAM ANDREW MOCK, PI
TREVOR DUNCAN RAY, SUSAN DEBORAH WHARAM, DONALD LEONARD NICHOLAS
PI CARDY
PC C12N15/09, C12Q1/68, C12N15/00
CC Description of Artificial Sequence: synthetic oligonucleotide
FT key Location/Qualifiers
FT source 1..12
FT /organism='Artificial Sequence'.
FEATURES source
1..12
Location/Qualifiers
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.1e+05; Length 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 1 ATCGTCAG 8
RESULT 3
AX019721
LOCUS AX019721 12 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9337806.
ACCESSION AX019721
VERSION AX019721.1 GI:10043561
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Cardy, D.L., Ray, T.D., Marsh, P., Assenberg, R., Weston, A., Mock, G.A. and Wharam, S.D.
TITLE Modified nucleic acid probes and uses thereof
JOURNAL Patent: WO 9337806-A 1 29-JUL-1999;
CARDY DONALD LEONARD NICHOLAS (GB); RAY TREVOR DUNCAN (GB); MARSH PETER (GB); ASSENBERG RENE (GB); CYTOCELL LTD (GB); WESTON ANTHONY (GB); MOCK GRAHAM ANDREW (GB); WHARAM SUSAN DEBORAH (GB)
FEATURES source
1..12
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic oligonucleotide"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.1e+05; Length 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 1 ATCGTCAG 8
RESULT 4
AX080675
LOCUS AX080675 12 bp DNA linear PAT 27-FEB-2001
DEFINITION Sequence 5 from Patent WO0109377.
ACCESSION AX080675

AX080675.1 GI:13169672
VERSION synthetic construct
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Lloyd, J.S., Weston, A. and Cardy, D.L.
TITLE Method for detecting nucleic acid target sequences involving in vitro transcription from an rna polymerase promoter
JOURNAL Patent: WO 0109377-A 5 08-FEB-2001;
CYTOCELL LIMITED (GB)
FEATURES source
1..12
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic oligonucleotide"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.1e+05; Length 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 1 ATCGTCAG 8
RESULT 5
AR011717/c
LOCUS AR011717 14 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 5 from patent US 5763169.
ACCESSION AR011717
VERSION AR011717.1 GI:3969707
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 14)
AUTHORS Sandhu, G.S. and Kline, B.C.
TITLE Nucleic acid probes for the detection and identification of fungi
JOURNAL Patent: US 5763169-A 5 09-JUN-1998;
FEATURES source
1..14
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1e+05; Length 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 11 ATCGTCAG 4
RESULT 6
AR075714/c
LOCUS AR075714 14 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 5 from patent US 5958693.
ACCESSION AR075714
VERSION AR075714.1 GI:10002460
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 14)
AUTHORS Sandhu, G.S. and Kline, B.C.
TITLE Extraction of DNA by boiling cells in an alkaline phenol/guanidine thiocyanate solution
JOURNAL Patent: US 5958693-A 5 28-SEP-1999;
FEATURES Location/Qualifiers


```
source
1. .14
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 11 ATCGTCAG 4

RESULT 7
LOCUS AR126298 14 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6180339.
ACCESSION AR126298
VERSION AR126298.1 GI:14112891
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Sandhu, G.S. and Kline, B.C.
TITLE Nucleic acid probes for the detection and identification of fungi
JOURNAL Patent: US 6180339-A 5 30-JAN-2001;
FEATURES
source
1. .14
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 11 ATCGTCAG 4

RESULT 8
LOCUS BD132864 14 bp DNA linear PAT 18-SEP-2002
DEFINITION Nucleic acid probes for the detection and identification of fungi.
ACCESSION BD132864
VERSION BD132864.1 GI:23227809
KEYWORDS JP 2002504817-A/5.
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Sandhu, G.S. and Kline, B.C.
TITLE Nucleic acid probes for the detection and identification of fungi
JOURNAL Patent: JP 2002504817-A 5 12-FEB-2002;
COMMENT BAYER CORP
PN JP 2002504817-A/5
PD 12-FEB-2002
PF 04-JUN-1998 JP 1999501953
PR 06-JUN-1997 US 08/871678
PI GURPREET S SANDHU, BRUCE C KLINE
PC C12Q1/68
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1. .14
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 4 ATCGTCAG 11

RESULT 10
LOCUS I79345/c 14 bp DNA linear PAT 10-JUN-1998
DEFINITION Sequence 5 from patent US 5707802.
ACCESSION I79345
VERSION I79345.1 GI:3207635
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 14)
AUTHORS Sandhu, G.S. and Kline, B.C.
TITLE Nucleic acid probes for the detection and identification of fungi
JOURNAL Patent: US 5707802-A 5 13-JAN-1998;
```

```

FEATURES
  source
    Location/Qualifiers
      1..14
      /organism="unknown"
      /mol_type="unassigned DNA"

ORIGIN

Query Match
  Best Local Similarity 100.0%; Score 8; DB 6; Length 14;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 11 ATCGTCAG 4

RESULT 11
LOCUS
  AX011314
  DEFINITION
    Sequence 18 from Patent WO9955908.
  ACCESSION
    AX011314
  VERSION
    AX011314.1 GI:9997864
  KEYWORDS
    .
  SOURCE
    Escherichia coli
  ORGANISM
    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
    Enterobacteriaceae; Escherichia.
  REFERENCE
    1
    AUTHORS
      Thierry,D., Frechon,D.T. and Laure,P.C.
    TITLE
      Nucleotide sequences for detecting enterohemorrhagic escherichia
      coli (ehc)
    JOURNAL
      PATENT: WO 9955908-A 18 04-NOV-1999;
      THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
      FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
  FEATURES
    source
      Location/Qualifiers
        1..14
        /organism="Escherichia coli"
        /mol_type="unassigned DNA"
        /db_xref="taxon:562"

ORIGIN

Query Match
  Best Local Similarity 100.0%; Score 8; DB 6; Length 14;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 4 ATCGTCAG 11

RESULT 12
LOCUS
  AR301299/c
  DEFINITION
    Sequence 61 from patent US 6538122.
  ACCESSION
    AR301299
  VERSION
    AR301299.1 GI:31689072
  KEYWORDS
    .
  SOURCE
    Unknown.
  ORGANISM
    Unclassified.
  REFERENCE
    1 (bases 1 to 15)
    AUTHORS
      Yokoyama,K., Nakamura,N., Miwa,T. and Seguro,K.
    TITLE
      Process for producing microbial transglutaminase
    JOURNAL
      Patent: US 6538122-A 61 25-MAR-2003;
  FEATURES
    source
      Location/Qualifiers
        1..15
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN

Query Match
  Best Local Similarity 100.0%; Score 8; DB 6; Length 15;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 13
LOCUS
  AX103974/c
  DEFINITION
    Sequence 166 from Patent WO0122972.
  ACCESSION
    AX103974
  VERSION
    AX103974.1 GI:13920171
  KEYWORDS
    .
  SOURCE
    synthetic construct
  ORGANISM
    other sequences; artificial sequences.
  REFERENCE
    1
    AUTHORS
      Krieg,A.M., Schetter,C. and Vollmer,J.C.
    TITLE
      Immunostimulatory nucleic acids
    JOURNAL
      Patent: WO 0122972-A 166 05-APR-2001;
      UNIVERSITY OF IOWA RESEARCH FOUNDATION (US); Coley Pharmaceutical
      GmbH (DE)
  FEATURES
    Location/Qualifiers
      source
        1..15
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"

ORIGIN

Query Match
  Best Local Similarity 100.0%; Score 8; DB 6; Length 15;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 14
LOCUS
  AX355608/c
  DEFINITION
    Sequence 636 from Patent WO0197843.
  ACCESSION
    AX355608
  VERSION
    AX355608.1 GI:18620276
  KEYWORDS
    .
  SOURCE
    synthetic construct
  ORGANISM
    other sequences; artificial sequences.
  REFERENCE
    1
    AUTHORS
      Weiner,G. and Hartmann,G.
    TITLE
      Methods for enhancing antibody-induced cell lysis and treating
      cancer
    JOURNAL
      Patent: WO 0197843-A 636 27-DEC-2001;
      UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
  FEATURES
    source
      Location/Qualifiers
        1..15
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="Synthetic oligonucleotide-phosphodiester backbone"

ORIGIN

Query Match
  Best Local Similarity 100.0%; Score 8; DB 6; Length 15;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 15
LOCUS
  AX547027/c

```

```

DEFINITION Sequence 166 from Patent WO02053141.
ACCESSION AX547027
VERSION AX547027.1 GI:25812171
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE Inhibition of angiogenesis by nucleic acids
JOURNAL Patent: WO 02053141-A 166 11-JUL-2002;
Coley Pharmaceutical Group, Inc. (US)
FEATURES
source
1. .15
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic Sequence"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 16
LOCUS BD064833 15 bp DNA linear PAT 27-AUG-2002
DEFINITION Method for detecting the extent of binding of transcriptional
regulatory protein to oligodNA.
ACCESSION BD064833
VERSION BD064833.1 GI:22610436
KEYWORDS JP 2001275678-A/45.
SOURCE
ORGANISM
REFERENCE
AUTHORS Kishimoto,T., Niwa,S., Mori,Y., Sachiyo, Mimaki, Fukushima,R. and
Nishikawa,K.
TITLE Method for detecting the extent of binding of transcriptional
regulatory protein to oligodNA
JOURNAL Patent: JP 2001275678-A 45 09-OCT-2001;
COMMENT SUMITOMO ELECTRIC INDUSTRIES LTD
OS Artificial Sequence
PN JP 2001275678-A/45
PD 09-OCT-2001
PF 31-MAR-2000 JP 2000096306
PI TOSHIHIKO KISHIMOTO,SHINICHIRO NIWA,YUKO MORI,SACHIYO PI
MIMAKI,REI FUKUSHIMA,
PI KAZUKO NISHIKAWA
PC C12N15/09,C12N5/10,C12Q1/00,C12Q1/68,C12N15/00,C12N5/00 CC
Synthetic DNA
FH Key Location/Qualifiers
FT source 1. .15
/organism="Artificial Sequence".
FEATURES
source
1. .15
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 12 ATCGTCAG 5

DEFINITION Sequence 166 from Patent WO02053141.
ACCESSION AX547027
VERSION AX547027.1 GI:25812171
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE Inhibition of angiogenesis by nucleic acids
JOURNAL Patent: WO 02053141-A 166 11-JUL-2002;
Coley Pharmaceutical Group, Inc. (US)
FEATURES
source
1. .15
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic Sequence"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 16
LOCUS BD064833 15 bp DNA linear PAT 27-AUG-2002
DEFINITION Method for detecting the extent of binding of transcriptional
regulatory protein to oligodNA.
ACCESSION BD064833
VERSION BD064833.1 GI:22610436
KEYWORDS JP 2001275678-A/45.
SOURCE
ORGANISM
REFERENCE
AUTHORS Kishimoto,T., Niwa,S., Mori,Y., Sachiyo, Mimaki, Fukushima,R. and
Nishikawa,K.
TITLE Method for detecting the extent of binding of transcriptional
regulatory protein to oligodNA
JOURNAL Patent: JP 2001275678-A 45 09-OCT-2001;
COMMENT SUMITOMO ELECTRIC INDUSTRIES LTD
OS Artificial Sequence
PN JP 2001275678-A/45
PD 09-OCT-2001
PF 31-MAR-2000 JP 2000096306
PI TOSHIHIKO KISHIMOTO,SHINICHIRO NIWA,YUKO MORI,SACHIYO PI
MIMAKI,REI FUKUSHIMA,
PI KAZUKO NISHIKAWA
PC C12N15/09,C12N5/10,C12Q1/00,C12Q1/68,C12N15/00,C12N5/00 CC
Synthetic DNA
FH Key Location/Qualifiers
FT source 1. .15
/organism="Artificial Sequence".
FEATURES
source
1. .15
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 12 ATCGTCAG 5

RESULT 17
LOCUS A72834 16 bp DNA linear PAT 15-OCT-1999
DEFINITION Sequence 14 from Patent WO9506126.
ACCESSION A72834
VERSION A72834.1 GI:6063860
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Hoekema,A. and Pen,J.
TITLE PRODUCTION OF TREHALOSE IN PLANTS
JOURNAL Patent: WO 9506126-A 14 02-MAR-1995;
MOGEN INT (NL); HOEKEMA ANDREAS (NL)
FEATURES
Location/Qualifiers
1. .16
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 2 ATCGTCAG 9

RESULT 18
LOCUS A72978 16 bp DNA linear PAT 15-OCT-1999
DEFINITION Sequence 14 from Patent WO9501446.
ACCESSION A72978
VERSION A72978.1 GI:6063909
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Hoekema,A. and Pen,J.
TITLE PRODUCTION OF TREHALOSE IN PLANTS
JOURNAL Patent: WO 9501446-A 14 12-JAN-1995;
MOGEN INT (NL); HOEKEMA ANDREAS (NL)
FEATURES
Location/Qualifiers
1. .16
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 2 ATCGTCAG 9

RESULT 19
LOCUS BD205229 16 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia
coli (EHEC).
ACCESSION BD205229
VERSION BD205229.1 GI:33014999
KEYWORDS JP 2002512813-A/19.
SOURCE

```

```

ORGANISM      unidentified
REFERENCE      1 (bases 1 to 16)
AUTHORS        Frechon,D.T.M., Laure,P.C. and Thierry,D.
TITLE          Nucleotide sequence for detecting enterohemorrhagic Escherichia
               coli (EHEC)
JOURNAL        Patent: JP 2002512813-A 19 08-MAY-2002;
COMMENT        BIORAD PASTER
OS             Unidentified
PN             JP 2002512813-A/19
PD             08-MAY-2002
PF             27-APR-1999 JP 2000546051
PR             28-APR-1998 FR 98/05329
PI             DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
PC             C12N9/08, C07K14/245, C12N1/21, C12N15/09, C12Q1/68, C12N15/00 CC
CC             Topology: Linear;
CC             Strandedness: Single;
CC             Nucleotide sequence for detecting enterohemorrhagic CC
CC             Escherichia coli
FH             Key
FT             Location/Qualifiers
               1..16
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"
FEATURES       source
               Location/Qualifiers
               1..16
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"
ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      5 ATCGTCAG 12

RESULT 20
AR329607/c
LOCUS          AR329607 16 bp RNA linear PAT 17-AUG-2003
DEFINITION     Sequence 7009 from patent US 6566127.
ACCESSION      AR329607
VERSION        AR329607.1 GI:33715415
KEYWORDS       .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 16)
AUTHORS        Pavco,P., McSwiggen,J.A., Stinchcomb,D.T. and Escobedo,J.
TITLE          Method and reagent for the treatment of diseases or conditions
               related to levels of vascular endothelial growth factor receptor
JOURNAL        Patent: US 6566127-A 7009 20-MAY-2003;
FEATURES       source
               1..16
               /organism="unknown"
               /mol_type="unassigned RNA"
ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      16 ATCGTCAG 9

RESULT 21
AR329608/c
LOCUS          AR329608 16 bp RNA linear PAT 17-AUG-2003
DEFINITION     Sequence 7010 from patent US 6566127.
ACCESSION      AR329608 1 GI:33715416
KEYWORDS       .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 16)
AUTHORS        Pavco,P., McSwiggen,J.A., Stinchcomb,D.T. and Escobedo,J.
TITLE          Method and reagent for the treatment of diseases or conditions
               related to levels of vascular endothelial growth factor receptor
JOURNAL        Patent: US 6566127-A 7010 20-MAY-2003;
FEATURES       source
               1..16
               /organism="unknown"
               /mol_type="unassigned RNA"
ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      5 ATCGTCAG 12

```

```

DEFINITION     Sequence 7010 from patent US 6566127.
ACCESSION      AR329608
VERSION        AR329608.1 GI:33715416
KEYWORDS       .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 16)
AUTHORS        Pavco,P., McSwiggen,J.A., Stinchcomb,D.T. and Escobedo,J.
TITLE          Method and reagent for the treatment of diseases or conditions
               related to levels of vascular endothelial growth factor receptor
JOURNAL        Patent: US 6566127-A 7010 20-MAY-2003;
FEATURES       source
               1..16
               /organism="unknown"
               /mol_type="unassigned RNA"
ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      13 ATCGTCAG 6

RESULT 22
AX011315
LOCUS          AX011315 16 bp DNA linear PAT 06-SEP-2000
DEFINITION     Sequence 19 from Patent WO9955908.
ACCESSION      AX011315
VERSION        AX011315.1 GI:9997865
KEYWORDS       .
SOURCE         Escherichia coli
               Escherichia coli
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Escherichia.
REFERENCE      1
AUTHORS        Thierry,D., Frechon,D.T. and Laure,P.C.
TITLE          Nucleotide sequences for detecting enterohemorrhagic escherichia
               coli (ehc)
JOURNAL        Patent: WO 9955908-A 19 04-NOV-1999;
               THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
               FRANCOISE CLAUDINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES       source
               1..16
               /organism="Escherichia coli"
               /mol_type="unassigned DNA"
               /db_xref="taxon:562"
ORIGIN
Query Match      100.0%; Score 8; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATCGTCAG 8
Db      5 ATCGTCAG 12

RESULT 23
BD093184/c
LOCUS          BD093184 16 bp DNA linear PAT 27-AUG-2002
DEFINITION     A gene coding a cyclic lipo peptide acylase and an expression
               thereof.
ACCESSION      BD093184
VERSION        BD093184.1 GI:22638772
KEYWORDS       WO 0102585-A/47.
SOURCE         synthetic construct
               other sequences; artificial sequences.
REFERENCE      1 (bases 1 to 16)

```

AUTHORS Shibata, T., Noguchi, Y. and Ymashita, M.
 TITLE A gene coading a cyclic lopoepetide acylase and an expression
 JOURNAL Patent: WO 0102585-A 47 11-JAN-2001;
 FUJISAWA PHARMACEUTICAL CO LTD, TAKASHI SHIBATA, YUJI NOGUCHI, MICHIO
 YMAASHITA

COMMENT

OS Artificial Sequence
 PN WO 0102585-A/47
 PD 11-JAN-2001
 PF 28-JUN-2000 WO 2000JP004285
 PR 02-JUL-1999 JP 99P 189644
 PI TAKASHI SHIBATA, YUJI NOGUCHI, MICHIO YMAASHITA
 PC C12N15/55, C12N1/21, C12N9/14
 CC Oligonucleotide designed to act as sequencing primer. FH Key

FEATURES

source
 1. .16
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 5.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8

Db 12 ATCGTCAG 5

RESULT 24

BD093185

LOCUS BD093185 16 bp DNA linear PAT 27-AUG-2002
 DEFINITION A gene coading a cyclic lopoepetide acylase and an expression
 thereof.

ACCESSION

BD093185

VERSION BD093185.1 GI:22638773

KEYWORDS WO 0102585-A/48.

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 16)

AUTHORS Shibata, T., Noguchi, Y. and Ymashita, M.

TITLE A gene coading a cyclic lopoepetide acylase and an expression

JOURNAL Patent: WO 0102585-A 48 11-JAN-2001;
 FUJISAWA PHARMACEUTICAL CO LTD, TAKASHI SHIBATA, YUJI NOGUCHI, MICHIO
 YMAASHITA

COMMENT

OS Artificial Sequence
 PN WO 0102585-A/48
 PD 11-JAN-2001
 PF 28-JUN-2000 WO 2000JP004285
 PR 02-JUL-1999 JP 99P 189644
 PI TAKASHI SHIBATA, YUJI NOGUCHI, MICHIO YMAASHITA
 PC C12N15/55, C12N1/21, C12N9/14
 CC Oligonucleotide designed to act as sequencing primer. FH Key

FEATURES

source
 1. .16
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 5.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8

Db 4 ATCGTCAG 11

RESULT 25

BD093185

LOCUS BD093185 17 bp DNA linear PAT 02-FEB-2004
 DEFINITION Sequence 8750 from Patent WO0192524.
 ACCESSION CQ624010
 VERSION CQ624010.1 GI:41674228
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 Gu, Y., Ji, Y., Penn, S.G., Hanzel, D.K., Rank, D.R., Chen, W. and
 Shannon, M.E.
 TITLE Myosin-like gene expressed in human heart and muscle
 JOURNAL Patent: WO 0192524-A 8750 06-DEC-2001;
 Aecomica, Inc. (US)
 FEATURES
 source
 1. .17
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 8; DB 6; Length 17;
 Best Local Similarity 100.0%; Pred. No. 5.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8

Db 10 ATCGTCAG 17

Search completed: March 11, 2005, 06:54:59

Job time : 46.8366 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 11.7034 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-1_COPY_400_407

Perfect score: 8

Sequence: 1 atcgtag 8

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870567 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	100.0	10	5	Aaf40099 Yeast NOR
C 2	8	100.0	10	5	Aaf39974 Yeast NOR
C 3	8	100.0	10	5	Aaf42756 Yeast NOR
C 4	8	100.0	12	2	Aax91001 Alternati
C 5	8	100.0	12	4	Aaf60612 T7 poly
C 6	8	100.0	14	2	Aat39175 Candida a
C 7	8	100.0	14	2	Aax05365 Nucleic
C 8	8	100.0	14	3	Aaz36118 Primer de
C 9	8	100.0	15	4	Aaf99041 Immunosti
C 10	8	100.0	15	6	Abs77682 Angiogene
C 11	8	100.0	15	6	ABL39211 Immunosti
C 12	8	100.0	15	9	ACD99473 Immunosti
C 13	8	100.0	15	9	ADB36543 Immunosti
C 14	8	100.0	15	12	ADH22637 Non-compl
C 15	8	100.0	16	2	Aaq84964 PCR prime
C 16	8	100.0	16	3	Aaz36119 Primer de
C 17	8	100.0	16	4	Aaf32294 Streptomy
C 18	8	100.0	16	4	Aaf32295 Streptomy
C 19	8	100.0	17	2	Aax71182 Human KDR
C 20	8	100.0	17	2	Aax71183 Human KDR

C	21	8	100.0	17	3	AAA70610 Sindbis-1
	22	8	100.0	17	3	AAA25898 Oestrogen
	23	8	100.0	17	3	AAA25897 Oestrogen
	24	8	100.0	17	6	ABN08761 Human GDM
	25	8	100.0	17	6	ABN08764 Human GDM
	26	8	100.0	17	6	ABN08766 Human GDM
	27	8	100.0	17	6	ABN08765 Human GDM
	28	8	100.0	17	6	ABN08767 Human GDM
	29	8	100.0	17	6	ABN08758 Human GDM
	30	8	100.0	17	6	ABN08760 Human GDM
	31	8	100.0	17	6	ABN08762 Human GDM
	32	8	100.0	17	6	ABN08763 Human GDM
	33	8	100.0	17	6	ABN08759 Human GDM
	34	8	100.0	17	6	AAS95508 Primer #2
	35	8	100.0	17	8	AAS53194 Candida a
	36	8	100.0	17	8	ABT37899 Tumour su
	37	8	100.0	17	8	ABT35225 Tumour su
	38	8	100.0	17	8	ABT35835 Tumour su
	39	8	100.0	17	8	ABT38114 Tumour su
	40	8	100.0	17	8	ACC66466 Murine ol
	41	8	100.0	17	8	ACC66260 Murine ol
	42	8	100.0	17	8	ACC66154 Murine ol
	43	8	100.0	17	8	ACC64969 Murine ol
	44	8	100.0	17	10	ADB39999 Tumour su
	45	8	100.0	17	10	ADI48451 Human tum

ALIGNMENTS

RESULT 1

Aaf40099/c

ID AAF40099 standard; DNA; 10 BP.

XX

AC AAF40099;

XX

DT 23-MAR-2001 (first entry)

XX

DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:6838.

XX

KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;

KW nor previously assigned open reading frame; nonannotated ORF; SAGE;

KW serial analysis of gene expression; antifungal; tag; identification;

KW linker; PCR primer; ds.

XX

OS Saccharomyces cerevisiae.

XX

PN WO200077214-A2.

XX

PD 21-DEC-2000.

XX

PF 14-JUN-2000; 2000WO-US016223.

XX

PR 16-JUN-1999; 99US-00335032.

XX

FA (UYJO) UNIV JOHNS HOPKINS.

XX

PI Veiculescu V, Vogelstein B, Kinzler K;

XX

DR WPI; 2001-061874/07.

XX

PT Yeast gene coding sequences comprising NORF genes with serial analysis of

XX gene expression (SAGE) tags, useful for studying, monitoring and

XX affecting phases of the cell cycle.

XX Example; Page 244; 419pp; English.

XX

CC The present invention describes an isolated DNA molecule comprising a

CC coding sequence of a yeast gene selected from a group of 745 NORF (not

CC previously assigned open reading frame; or nonannotated ORF) genes

CC comprising a SAGE (serial analysis of gene expression) tag. Also

CC described are: (1) a method (M1) of using NORF genes to affect the cell

CC cycle comprising administering a NORF gene whose expression varies by at

CC least 10% between any two phases of the cell cycle selected from log
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
 CC antifungal drugs comprising: (a) contacting a test substance with a yeast
 CC cell; and (b) monitoring expression of a NORF gene whose expression
 CC varies as in M1, where a test substance which modifies the expression of
 CC the yeast gene is a candidate antifungal drug; (3) a method (M3) for
 CC identifying human genes which are involved in cell cycle progression
 CC comprising contacting human DNA with a probe which comprises at least 10
 CC contiguous nucleotides of a NORF gene whose expression varies as in M1;
 CC and (4) a method (M4) for identifying a candidate drug as a member of a
 CC class of drugs having a characteristic effect on gene expression in a
 CC yeast cell comprising contacting a yeast cell with a candidate drug and
 CC monitoring expression in the yeast cell of at least 1 NORF gene whose
 CC expression is affected by the class of drugs. The NORF genes may be used
 CC to study, monitor and affect phases of the cell cycle, the differentially
 CC expressed genes may be used to identify candidate drugs which affect the cell
 CC cycle and for identification of antifungal drugs. AAF33268 to AAF44064
 CC represent SAGE tags used in the exemplification of the present invention.
 CC AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE
 CC method, in the exemplification of the present invention
 XX
 SQ Sequence 10 BP; 3 A; 2 C; 2 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.5e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 10 ATCGTCAG 3

RESULT 2

AAF39974
 ID AAF39974 standard; DNA; 10 BP.

AC AAF39974;

XX 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:6713.

XX Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
 KW serial analysis of gene expression; antifungal; tag; identification;
 KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

OS WO200077214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US016223.

XX 16-JUN-1999; 99US-00335032.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI; 2001-061874/07.

XX Yeast gene coding sequences comprising NORF genes with serial analysis of
 PT gene expression (SAGE) tags, useful for studying, monitoring and
 PT affecting phases of the cell cycle.

XX Example; Page 239; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a
 CC coding sequence of a yeast gene selected from a group of 745 NORF (not
 CC previously assigned open reading frame; or nonannotated ORF) genes

CC comprising a SAGE (serial analysis of gene expression) tag. Also
 CC described are: (1) a method (M1) of using NORF genes to affect the cell
 CC cycle comprising administering a NORF gene whose expression varies by at
 CC least 10% between any two phases of the cell cycle selected from log
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
 CC antifungal drugs comprising: (a) contacting a test substance with a yeast
 CC cell; and (b) monitoring expression of a NORF gene whose expression
 CC varies as in M1, where a test substance which modifies the expression of
 CC the yeast gene is a candidate antifungal drug; (3) a method (M3) for
 CC identifying human genes which are involved in cell cycle progression
 CC comprising contacting human DNA with a probe which comprises at least 10
 CC contiguous nucleotides of a NORF gene whose expression varies as in M1;
 CC and (4) a method (M4) for identifying a candidate drug as a member of a
 CC class of drugs having a characteristic effect on gene expression in a
 CC yeast cell comprising contacting a yeast cell with a candidate drug and
 CC monitoring expression in the yeast cell of at least 1 NORF gene whose
 CC expression is affected by the class of drugs. The NORF genes may be used
 CC to study, monitor and affect phases of the cell cycle, the differentially
 CC expressed genes may be used as markers of phases of the cell cycle. The
 CC methods may be used to identify candidate drugs which affect the cell
 CC cycle and for identification of antifungal drugs. AAF33268 to AAF44064
 CC represent SAGE tags used in the exemplification of the present invention.
 CC AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE
 CC method, in the exemplification of the present invention
 XX
 SQ Sequence 10 BP; 3 A; 2 C; 2 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.5e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 3 ATCGTCAG 10

RESULT 3

AAF42756
 ID AAF42756 standard; DNA; 10 BP.

XX AAF42756;

XX 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:10895.

XX Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;
 KW nor previously assigned open reading frame; nonannotated ORF; SAGE;
 KW serial analysis of gene expression; antifungal; tag; identification;
 KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO200077214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US016223.

XX 16-JUN-1999; 99US-00335032.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI; 2001-061874/07.

XX Yeast gene coding sequences comprising NORF genes with serial analysis of
 PT gene expression (SAGE) tags, useful for studying, monitoring and
 PT affecting phases of the cell cycle.

XX Example; Page 339; 419pp; English.

CC The present invention describes an isolated DNA molecule comprising a
 CC coding sequence of a yeast gene selected from a group of 745 NORF genes
 CC previously assigned open reading frame; or nonannotated ORF) genes
 CC comprising a SAGE (serial analysis of gene expression) tag. Also
 CC described are: (1) a method (M1) of using NORF genes to affect the cell
 CC cycle comprising administering a NORF gene whose expression varies by at
 CC least 10% between any two phases of the cell cycle selected from log
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate
 CC antifungal drugs comprising: (a) contacting a test substance with a yeast
 CC cell; and (b) monitoring expression of a NORF gene whose expression
 CC varies as in M1, where a test substance which modifies the expression of
 CC the yeast gene is a candidate antifungal drug; (3) a method (M3) for
 CC identifying human genes which are involved in cell cycle progression
 CC comprising contacting human DNA with a probe which comprises at least 10
 CC contiguous nucleotides of a NORF gene whose expression varies as in M1;
 CC and (4) a method (M4) for identifying a candidate drug as a member of a
 CC class of drugs having a characteristic effect on gene expression in a
 CC yeast cell comprising contacting a yeast cell with a candidate drug and
 CC monitoring expression in the yeast cell of at least 1 NORF gene whose
 CC expression is affected by the class of drugs. The NORF genes may be used
 CC to study, monitor and affect phases of the cell cycle, the differentially
 CC expressed genes may be used as markers of phases of the cell cycle. The
 CC methods may be used to identify candidate drugs which affect the cell
 CC cycle and for identification of antifungal drugs. AAF33268 to AAF44064
 CC represent SAGE tags used in the exemplification of the present invention.
 CC AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE
 CC method, in the exemplification of the present invention

XX SQ Sequence 10 BP; 2 A; 3 C; 2 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4.5e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 3 ATCGTCAG 10

RESULT 4
 AAX91001
 ID AAX91001 standard; DNA; 12 BP.

XX AC AAX91001;

XX DT 08-OCT-1999 (first entry)

XX DE Alternative template sequence for T7 polymerase.

XX KW Nucleic acid detection; target; hybridisation; pathogen; human;
 XX KW point mutation detection; template; T7 polymerase; ss.

XX OS Synthetic.

XX PN WO9937806-A2.

XX PD 29-JUL-1999.

XX PF 26-JAN-1999; 99WO-GB000269.

XX PR 27-JAN-1998; 98GB-00001628.

XX PR 29-APR-1998; 98GB-00009014.

XX PA (CYTO-) CYTOCELL LTD.

XX PI Weston A. Assenberg R, Marsh P, Mock GA, Ray TD, Wharam SD;
 XX PI Cardy DLN;

XX XX WPI; 1998-479057/40.

XX DR Method for detection of nucleic acid target sequences, useful for
 XX PT detecting point mutations in a sequence of interest.

XX

PS Disclosure; Page 8; 81pp; English.

XX The invention provides a method for detection of nucleic acid target
 CC sequences in a sample. The method comprises (a) contacting the sample
 CC with first and second nucleic acid probes, where the first probe
 CC comprises a portion complementary to the sequence of interest and so
 CC capable of hybridising with it, and a portion non-complementary to the
 CC sequence of interest, and where the second probe comprises a portion
 CC complementary to the sequence of interest and is capable of hybridising
 CC with it, and a portion non-complementary to the sequence of interest but
 CC complementary to that portion of the first probe which is non-
 CC complementary to the sequence of interest, such that the first and second
 CC probes are capable of hybridising to the sequence of interest in adjacent
 CC or substantially adjacent manner, so as to allow complementary portions
 CC of the first and second probes to hybridise to each other; (b) causing
 CC extension of the first probe with a nucleic acid polymerase, using the
 CC second probe as a template; and (c) detecting directly or indirectly the
 CC extension of the first probe, so as to indicate the presence of the
 CC sequence of interest; characterised in that the first and/or second probe
 CC comprises a destabilizing moiety which cannot base pair with the
 CC reciprocal probe, therefore preventing hybridisation of the first and
 CC second probes in the absence of the sequence of interest. The method and
 CC probes can be used to detect a sequence from a pathogen or a human or
 CC allele. They can also detect point mutations in a sequence of interest.
 CC Sequences AAX91001-10 represent alternative template sequences for T7
 CC polymerase used in the course of the invention

XX SQ Sequence 12 BP; 2 A; 5 C; 2 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 4.5e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 1 ATCGTCAG 8

RESULT 5
 AAF60612

ID AAF60612 standard; DNA; 12 BP.

XX AC AAF60612;

XX DT 11-SEP-2003 (revised)

XX DT 06-AUG-2003 (revised)

XX DT 27-APR-2001 (first entry)

XX DE T7 polymerase promoter +12 region sequence #1.

XX KW T7 RNA polymerase promoter; probe; ds.

XX OS Enterobacteria phage T7.

XX PN WO200109377-A1.

XX PD 08-FEB-2001.

XX PF 31-JUL-2000; 2000WO-GB002962.

XX PR 29-JUL-1999; 99GB-00017813.

XX PR 17-AUG-1999; 99US-0149176P.

XX XX (CYTO-) CYTOCELL LTD.

XX XX Lloyd JS, Weston A, Cardy DLN;

XX XX WPI; 2001-182976/18.

XX PT New complex formed by a hybridization reaction, useful for detecting a
 XX PT nucleic acid sequence of interest in sample, comprises the target nucleic
 XX PT acid molecule and 2 or 3 probes.

PS Disclosure; Page 7; 58pp; English.

XX The present invention relates to a complex formed by the hybridisation of

CC a target nucleic acid (e.g. a bacteriophage RNA polymerase promoter

CC sequence) and probes. The probes are useful in an assay for detecting the

CC presence of a nucleic acid sequence of interest, in a sample. The present

CC sequence is a +12 region of 17 RNA polymerase promoter which was used in

CC the present invention. (Updated on 06-AUG-2003 to correct OS field.)

CC (Updated on 11-SEP-2003 to standardise OS field)

XX

SQ Sequence 12 BP; 2 A; 5 C; 2 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 4.5e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8

Db 1 ATCGTCAG 8

RESULT 6

AAT39175/c

ID AAT39175 standard; DNA; 14 BP.

XX

AC AAT39175;

XX

DT 25-FEB-1997 (first entry)

XX

DE Candida albicans 28S rRNA-specific probe.

XX

XX Species-specific; fungus; fungal; discrimination; detection; distinguish;

KW Acromonium; Aspergillus; Beauveria; Bipolaris; Blastoschizomyces;

KW Blastomyces; Candida; Chrysosporium; Cladosporium; Coccidioides;

KW Cryptococcus; Curvularia; Fusarium; Filobasidium; Geotrichum; Histoplasma;

KW Mucor; Malbranchea; Paecilomyces; Penicillium; Pseudallescheria;

KW Rhizopus; Sporothrix; Scopulariopsis; Saccharomyces; Trichosporon; probe;

XX ss.

XX Synthetic.

XX

XX WO9621741-A1.

XX

PD 18-JUL-1996.

XX

XX 12-JAN-1996; 96WO-IB000026.

XX

PR 13-JAN-1995; 95US-00373127.

PR 05-MAY-1995; 95US-00435684.

XX

XX (CIBA) CIBA CORNING DIAGNOSTICS CORP.

XX

XX Sandhu GS, Kline BC;

XX

DR WPI; 1996-342293/34.

XX

XX New oligonucleotide probes specific for 28S subunit of fungi - used to

PT identify particular fungal species, esp. after PCR amplification with

PT universal primers.

XX

PS Claim 1; Page 46; 108pp; English.

XX

XX AAT39173-T39244 are species-specific probes for the 28S rRNA subunit of

CC various types of fungi. The probes will hybridise only to 28S rRNA

CC sequences of one specific fungus species (or in a few cases a particular

CC genus). The probes are used to detect fungi that cause disease in humans

CC and animals or spoilage of foods and beverages. The individual probes are

CC able to identify one of the following genera and species: Acromonium sp.,

CC Aspergillus clavatus, A. flavus, A. fumigatus, A. glaucus, A. nidulans,

CC A. niger, A. ochraceus, A. terreus, A. unguis or A. ustus, Beauveria sp.,

CC Bipolaris sp., Blastoschizomyces sp., Blastomyces dermatitidis, Candida

CC albicans, C. glabrata, C. guilliermondii, C. kefyr, C. krusei, C.

CC lusitanae, C. parapsilosis, C. tropicalis, Chrysosporium sp.,

CC

CC Cladosporium sp., Coccidioides immitis, Cryptococcus neoformans var.

CC gattii serotype A or B, var. bacillispora serotype C or D, C. laurentii,

CC C. terreus, C. curvularia sp., Fusarium sp., Filobasidium capsuligenum, F.

CC unguiculatum, Geotrichum sp., Histoplasma capsulatum, Malbranchea sp.,

CC Mucor sp., Paecilomyces sp., Penicillium sp., Pseudallescheria boydii,

CC Rhizopus sp., Sporothrix schenckii, Scopulariopsis brevicaulis, S.

CC brumpti, Saccharomyces cerevisiae or Trichosporon beigellii. Those species

CC written in abbreviated form have the same genus name as the closest

CC previous species written out in full, e.g. C. laurentii is of the genus

CC Cryptococcus

XX

SQ Sequence 14 BP; 2 A; 5 C; 3 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.6e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8

Db 11 ATCGTCAG 4

RESULT 7

AAX05365/c

ID AAX05365 standard; DNA; 14 BP.

XX

AC AAX05365;

XX

DT 22-APR-1999 (first entry)

XX

DE Nucleic acid probe for Candida albicans.

XX

XX Fungus; disease; food spoilage; beverage; polymerase chain reaction; PCR;

KW nucleic acid hybridisation; Paracoccidioides brasiliensis;

KW Pneumocystis carinii; probe; ss.

XX

OS Synthetic.

OS Candida albicans.

XX

XX WO9855649-A1.

XX

PD 10-DEC-1998.

XX

XX 04-JUN-1998; 98WO-IB000865.

XX

PR 06-JUN-1997; 97US-00871678.

XX

XX (CHIR) CHIRON DIAGNOSTICS CORP.

XX

XX Sandhu GS, Kline BC;

XX

DR WPI; 1999-080835/07.

XX

XX Oligonucleotide primers and probes specific for fungal species - causing

PT diseases in animals or human and for food- and beverage spoiling fungi.

XX

PS Example 1; Page 56; 124pp; English.

XX

XX The invention relates to nucleic acid probes and primers for detecting a

CC wide variety of fungal species which cause disease in animals and humans

CC and also cause spoilage in food and beverage. These probes can detect

CC rRNA, rDNA or polymerase chain reaction (PCR) products from a majority of

CC fungi in clinical, environmental or food samples. Nucleic acid

CC hybridisation assay probes for detecting Paracoccidioides brasiliensis

CC (AAX05435) and Pneumocystis carinii (AAX05436) are specifically claimed.

CC Sequences AAX05363-383 represent probes for different fungal species (see

CC descriptor line for the specific fungal species) that were tested for

CC hybridisation specificity

XX

SQ Sequence 14 BP; 2 A; 5 C; 3 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. No. 4.6e+04;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 11 ATCGTCAG 4

RESULT 8
 AAZ36118
 ID AAZ36118 standard; DNA; 14 BP.
 XX AC
 XX AC
 XX XX
 XX XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Primer derived from a nucleic acid sequence specific to EHEC.
 XX
 KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 KW PCR primer; probe; ss.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX WO9955908-A2.
 XX
 XX 04-NOV-1999.
 XX
 XX 27-APR-1999; 99WO-FR001000.
 XX
 XX 28-APR-1998; 98FR-00005329.
 XX
 XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.
 XX
 PI Frechon DTM, Laure FC, Thierry D;
 XX
 XX WPI; 2000-013443/01.
 XX
 XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.
 XX
 PS Claim 5; Page 27; 48pp; French.
 XX
 CC AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies
 XX
 SQ Sequence 14 BP; 2 A; 3 C; 5 G; 4 T; 0 U; 0 Other;
 Query Match 100.0%; Score 8; DB 3; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 4 ATCGTCAG 11

RESULT 9
 AAF99041/c
 ID AAF99041 standard; DNA; 15 BP.
 XX

AC AAF99041;
 XX 12-JUN-2001 (first entry)
 DT
 XX
 DE Immunostimulatory nucleic acid #157.
 XX
 KW Vaccine; cytostatic; virucidal; bactericidal; fungicidal; anti-parasitic;
 KW immunostimulatory; tumour; viral infection; bacterial infection;
 KW fungal infection; parasitic infection; cancer; asthma;
 KW infectious disease; allergy; immune deficiency; phosphorothioate; ss.
 XX
 OS Synthetic.
 OS
 XX WO200122972-A2.
 XX
 XX 05-APR-2001.
 XX
 XX 25-SEP-2000; 2000WO-US026383.
 XX
 XX 25-SEP-1999; 99US-0156113P.
 PR
 XX 27-SEP-1999; 99US-0156135P.
 PR
 XX 23-AUG-2000; 2000US-0227436P.
 PR
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (COLE-) COLEY PHARM GMBH.
 PA
 XX Krieg AM, Schetter C, Vollmer J;
 PI
 XX WPI; 2001-273485/28.
 DR
 XX
 XX Vaccinating against tumors, infectious diseases, allergies and asthma
 PT using immunostimulatory Py-rich and TG nucleic acids.
 XX
 PS Claim 101; Page 41; 338pp; English.
 XX
 CC The present invention relates to a method for stimulating an immune
 CC response. The method comprises administering an immunostimulatory nucleic
 CC acid to a non-rodent subject in sufficient quantity to stimulate an
 CC immune response. The present sequence is one such immunostimulatory
 CC nucleic acid. The immunostimulatory nucleic acids can be pyrimidine rich
 CC (py-rich) or thymidine (T) rich. The method is used to vaccinate subjects
 CC against tumour antigens, viral antigens (e.g. herpesviridae, retroviridae
 CC and/or orthomyxoviridae), bacterial antigens (e.g. toxoplasma,
 CC haemophilus, campylobacter, clostridium, Escherichia coli and/or
 CC staphylococcus), fungal antigens and/or parasitic antigens. The method is
 CC also useful for preventing cancer, asthma, infectious disease, allergy or
 CC immune deficiency. The present sequence can also be used to redirect a
 CC Th2 to a Th1 immune response and to activate immune cells. Note: the
 CC present sequence may have a phosphorothioate backbone
 XX
 SQ Sequence 15 BP; 4 A; 3 C; 4 G; 4 T; 0 U; 0 Other;
 Query Match 100.0%; Score 8; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 10 ATCGTCAG 3

RESULT 10
 ABS77682/c
 ID ABS77682 standard; DNA; 15 BP.
 XX
 XX ABS77682;
 AC
 XX
 DT 13-DEC-2002 (first entry)
 XX
 DE Angiogenesis inhibitory oligonucleotide #166.
 XX
 XX Angiogenesis inhibitor; ss; angiogenesis; solid tumour growth;
 KW tumour metastasis; precancerous lesion; rheumatoid arthritis; psoriasis;

KW diabetic retinopathy; retinopathy of prematurity; macular degeneration;
KW corneal graft rejection; neovascular glaucoma; retrolental fibroplasia;
KW rubeosis; Osher-Webber Syndrome; myocardial angiogenesis;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; wound granulation; intestinal adhesion; atherosclerosis;
KW scleroderma; hypertrophic scar.
XX
OS Synthetic.
XX
PN WO200253141-A2.
XX
XX 11-JUL-2002.
XX
XX 14-DEC-2001; 2001WO-US048458.
XX
XX 14-DEC-2000; 2000US-0255534P.
XX
XX (COLE-) COLEY PHARM GROUP INC.
XX
XX Bratzler RL;
XX
XX WPI; 2002-566690/60.
XX
XX Inhibiting angiogenesis in a subject, involves administering at least one
XX antiangiogenic nucleic acid molecule to the subject.
XX
XX Claim 2; Page 22; 276pp; English.
XX
XX The invention relates to inhibiting angiogenesis in a subject, comprising
XX administering at least one antiangiogenic nucleic acid molecule. Also
XX included is a kit comprising a first container housing the antiangiogenic
XX nucleic acids, and instructions for administering them to a subject
XX having a condition characterised by unwanted angiogenesis. The method is
XX useful for inhibiting angiogenesis associated with solid tumour growth,
XX tumour metastasis, precancerous lesion, rheumatoid arthritis, psoriasis,
XX diabetic retinopathy, retinopathy of prematurity, macular degeneration,
XX corneal graft rejection, neovascular glaucoma, retrolental fibroplasia,
XX rubeosis, Osher-Webber Syndrome, myocardial angiogenesis, plaque
XX neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
XX wound granulation, intestinal adhesions, atherosclerosis, scleroderma and
XX hypertrophic scars. The present sequence is an antiangiogenic nucleic
XX acid of the invention
SQ Sequence 15 BP; 4 A; 3 C; 4 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 8; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3
RESULT 11
ABL39211/c
ID ABL39211 standard; DNA; 15 BP.
XX
XX ABL39211;
AC
XX
XX 16-APR-2002 (first entry)
XX
XX Immunostimulatory nucleic acid SEQ ID NO: 636.
XX
XX Antibody-induced cell lysis; cancer; immunostimulatory; CD20;
KW angiogenesis; metastasis; cytostatic; ss.
XX
XX Synthetic.
OS
XX WO200197843-A2.
PN
XX 27-DEC-2001.
PD
XX

PF 22-JUN-2001; 2001WO-US020154.
XX
XX 22-JUN-2000; 2000US-0213346P.
XX
XX (IOWA) UNIV IOWA RES FOUND.
XX
XX Weiner G, Hartmann G;
PI
XX WPI; 2002-154611/20.
DR
XX
XX Treating or preventing cancer, such as basal cell carcinoma, comprises
XX administering immunostimulatory nucleic acids that induce expression of
XX cell surface antigens and antibodies to a subject having or at risk of
XX developing cancer.
XX
XX Disclosure; Page 257; 312pp; English.
XX
XX The present invention relates to methods for treating or preventing
XX cancer, involving administering to a subject having or at risk of
XX developing cancer immunostimulatory nucleic acids that induce expression
XX of cell surface antigens and antibodies. The methods are useful for
XX treating or preventing cancer such as basal cell carcinoma, bladder
XX cancer, bone cancer, brain and central nervous system (CNS) cancer,
XX breast cancer, cervical cancer, colon and rectum cancer, connective
XX tissue cancer, oesophageal cancer, eye cancer, kidney cancer, larynx
XX cancer, leukaemia, liver cancer, lung cancer, Hodgkin's lymphoma, non-
XX Hodgkin's lymphoma, melanoma, myeloma, oral cavity cancer, ovarian
XX cancer, pancreatic cancer, prostate cancer, rhabdomyosarcoma, skin
XX cancer, stomach cancer, testicular cancer, and uterine cancer. The
XX present sequence is an immunostimulatory oligonucleotide described in the
XX exemplification of the invention
SQ Sequence 15 BP; 4 A; 3 C; 4 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 8; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3
RESULT 12
ACD99473/c
ID ACD99473 standard; DNA; 15 BP.
XX
XX ACD99473;
AC
XX
XX 25-SEP-2003 (first entry)
DT
XX
XX Immunostimulatory nucleic acid #159.
XX
XX Immunostimulatory; antiinflammatory; dermatological; antipsoriatic;
KW antiulcer; gene therapy; vaccine; non-allergic inflammatory disease;
KW psoriasis; eczema; allergic contact dermatitis; latex dermatitis;
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease; ss.
XX
XX Synthetic.
OS
XX US2003050268-A1.
PN
XX 13-MAR-2003.
PD
XX
XX 29-MAR-2002; 2002US-00112653.
PF
XX
XX 29-MAR-2001; 2001US-0279642P.
PR
XX
XX (KRIE/) KRIEG A M.
XX PA
XX (BERG/) BERG D J.
XX PA
XX Krieg AM, Berg DJ;
PI
XX

DR WPI; 2003-521815/49.
XX Treating non-allergic inflammatory diseases, such as psoriasis, eczema,
PT allergic contact dermatitis, latex dermatitis or inflammatory bowel
PT disease by administering an immunostimulatory nucleic acid.
XX
XX
XX Disclosure; Page 12; 229pp; English.
XX
XX The invention describes a method of treating non-allergic inflammatory
CC disease comprising administering to a subject having or at risk of
CC developing a non-allergic inflammatory disease an immunostimulatory
CC nucleic acid for prevention or treatment of the disease. The method is
CC useful for treating non-allergic inflammatory diseases, such as
CC psoriasis, eczema, allergic contact dermatitis, latex dermatitis or
CC inflammatory bowel disease e.g., ulcerative colitis or Crohn's disease.
CC This sequence represents an immunostimulatory nucleic acid
XX
XX Sequence 15 BP; 4 A; 3 C; 4 G; 4 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 8; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGTCAG 8
| | | | |
DB 10 ATCGTCAG 3

RESULT 13
ADB36543/C
ID ADB36543 standard; DNA; 15 BP.
XX
XX ADB36543;
XX
XX 04-DEC-2003 (first entry)
XX
XX Immunostimulatory nucleic acid #157.
XX
XX ds; allergy; asthma; poly-G nucleic acid; aerosol formulation;
XX hypo-responsive subject; immunostimulatory.
XX Synthetic.
XX
XX US2003087848-A1.
XX
XX 08-MAY-2003.
XX
XX 02-FEB-2001; 2001US-00776479.
XX
XX 03-FEB-2000; 2000US-0179991P.
XX
XX (BRAT/) BRATZLER R L.
XX (PETE/) PETERSEN D M.
XX (FOUR/) FOURON Y.
XX
XX Bratzler RL, Petersen DM, Fouron Y;
XX WPI; 2003-657977/62.
XX
XX Treating and/or preventing allergy or asthma using an immunostimulatory
XX nucleic acid alone or in combination with an asthma/allergy medicament.
XX Disclosure; Page 7; 221pp; English.
XX
XX The invention relates to a method of treating or preventing allergy or
XX asthma which comprises administering to a subject a poly-G nucleic acid
XX in an aerosol formulation. The methods and compositions of the present
XX invention are useful for diagnosing and/or treating asthma and allergy
XX especially in a hypo-responsive subject. The present sequence represents
XX an immunostimulatory nucleic acid of the invention.
XX
XX Sequence 15 BP; 4 A; 3 C; 4 G; 4 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 8; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGTCAG 8
| | | | |
DB 10 ATCGTCAG 3

RESULT 14
ADH22637/C
ID ADH22637 standard; DNA; 15 BP.
XX
XX ADH22637;
XX
XX 11-MAR-2004 (first entry)
XX
XX Non-complementary DNA oligo used as a control.
XX
XX light harvesting multichromophore; sensor; peptide nucleic acid; PNA;
XX polycationic multichromophore; PNA-C*; fluorescein; ss.
XX Unidentified.
XX
XX WO2004001379-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019678.
XX
XX 20-JUN-2002; 2002US-0390524P.
XX
XX 26-AUG-2002; 2002US-0406266P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Bazan GC, Gaylord BS;
XX
XX WPI; 2004-142830/14.
XX
XX Assay method, by contacting sample e.g. blood, urine with sensor peptide
XX nucleic acid which has signaling chromophore to absorb energy from
XX excited multichromophore and emit light in presence of target
XX polynucleotide.
XX
XX Example 2; Page 21; 34pp; English.
XX
XX This invention relates to a novel method for the detection and analysis
XX of polynucleotides using light harvesting multichromophores.
XX Specifically, it refers to a method for assaying a target polynucleotide
XX that comprises contacting this target DNA with a complementary sensor
XX peptide nucleic acid (PNA) and a polycationic multichromophore. As such,
XX the present invention describes a sensor PNA that is conjugated to a
XX signalling chromophore which absorbs energy from the excited
XX multichromophore emitting light in the presence of the target DNA
XX molecule. Accordingly, this method is useful for assaying target nucleic
XX acids in samples such as blood, urine, milk, buccal, vaginal or rectal
XX swabs and aspirate. Furthermore, it allows the analysis of a target
XX polynucleotide that occurs naturally in a sample and by using multiple
XX different sensor PNAs, a plurality of polynucleotides can be detected and
XX assayed simultaneously. This oligonucleotide sequence is the non-
XX complementary DNA oligo that is used as a control in an exemplification
XX of the invention.
XX
XX Sequence 15 BP; 5 A; 3 C; 4 G; 3 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 8; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGTCAG 8
| | | | |
DB 9 ATCGTCAG 2

PT Gene encoding cyclic lipopeptide acylase genetically engineered to give
 PT vectors and transformants for expression of protein with comparable
 PT acylase activity in shorter culture time on large scale.
 XX Example 1; Page 24; 73pp; Japanese.

CC The present invention describes a Streptomyces sp. cyclic lipopeptide
 CC acylase. The cyclic lipopeptide acylase gene and its expressed cyclic
 CC lipopeptide acylase are useful in deacylation of the amino group in the
 CC acylamino group of a side-chain in a cyclic lipopeptide substance. Cyclic
 CC lipopeptide acylases are obtainable by genetic modification, have
 CC comparable acylase activity to the parent and can be produced in shorter
 CC culture time on large scale. The present sequence represents a sequencing
 CC primer for the Streptomyces sp. cyclic lipopeptide acylase, which is used
 CC in an example from the present invention

XX Sequence 16 BP; 3 A; 5 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 4; Length 16;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 12 ATCGTCAG 5

RESULT 18

AAF32295
 ID AAF32295 standard; DNA; 16 BP.

XX AAF32295;

DT 17-APR-2001 (first entry)

XX Streptomyces sp. cyclic lipopeptide acylase sequencing primer AC40.

XX Streptomyces; cyclic lipopeptide acylase; acylase; deacylation;
 KW acylamino group; sequencing primer; ss.

OS Streptomyces sp.

XX WO200102585-A1.

XX 11-JAN-2001.

XX 28-JUN-2000; 2000WO-JP004285.

XX 02-JUL-1999; 95JP-00189644.

XX (FUJII) FUJISAWA PHARM CO LTD.

XX Shibata T, Noguchi Y, Yamashita M;

XX WPI; 2001-123114/13.

XX Gene encoding cyclic lipopeptide acylase genetically engineered to give
 PT vectors and transformants for expression of protein with comparable
 PT acylase activity in shorter culture time on large scale.

XX Example 1; Page 24; 73pp; Japanese.

CC The present invention describes a Streptomyces sp. cyclic lipopeptide
 CC acylase. The cyclic lipopeptide acylase gene and its expressed cyclic
 CC lipopeptide acylase are useful in deacylation of the amino group in the
 CC acylamino group of a side-chain in a cyclic lipopeptide substance. Cyclic
 CC lipopeptide acylases are obtainable by genetic modification, have
 CC comparable acylase activity to the parent and can be produced in shorter
 CC culture time on large scale. The present sequence represents a sequencing
 CC primer for the Streptomyces sp. cyclic lipopeptide acylase, which is used
 CC in an example from the present invention

XX Sequence 16 BP; 3 A; 5 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 4; Length 16;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 4 ATCGTCAG 11

RESULT 19

AAAX71182/c

ID AAAX71182 standard; RNA; 17 BP.

XX AAAX71182;

XX 28-JUL-1999 (first entry)

XX Human KDR VEGF receptor hammerhead ribozyme substrate #194.

XX Vascular endothelial growth factor receptor; VEGF receptor; flt-1; flk-1;
 KW KDR; hammerhead ribozyme; hairpin ribozyme; cleavage;
 KW tumour angiogenesis; psoriasis; rheumatoid arthritis; ocular disease;
 KW fms-like tyrosine kinase 1; kinase insert domain containing receptor;
 KW foetal liver kinase 1; ss.

OS Homo sapiens.

XX WO9715662-A2.

XX 01-MAY-1997.

XX 25-OCT-1996; 96WO-US017480.

XX 26-OCT-1995; 95US-0005974P.

XX 11-JAN-1996; 96US-00584040.

XX (RIBO-) RIBOZYME PHARM INC.

XX (CHIR) CHIRON CORP.

XX Pavco P, Mcawiggen J, Stinchcomb D, Escobedo J;

XX WPI; 1997-259017/23.

XX Nucleic acid molecule modulating VEGF receptor(s) gene expression or mRNA
 PT stability - useful for treating e.g. tumour angiogenesis, psoriasis,
 PT rheumatoid arthritis, etc., in a human patient.

XX Claim 4; Page 103; 218pp; English.

XX The present invention describes nucleic acid molecules which modulate the
 CC synthesis, expression and/or stability of a mRNA encoding 1 or more
 CC receptors of vascular endothelial growth factor (VEGF). A patient
 CC (preferably human) having a condition associated with the level of the
 CC fms-like tyrosine kinase 1 (flt-1), kinase insert domain containing
 CC receptor (KDR) and/or foetal liver kinase 1 (flk-1) (e.g. tumour
 CC angiogenesis, ocular diseases, psoriasis and rheumatoid arthritis) can be
 CC treated by administering the nucleic acid molecule or the expression
 CC vector to the patient. AAAX7275 to AAAX7572 represent specific examples
 CC of nucleic acid molecules from the present invention

XX Sequence 17 BP; 4 A; 3 C; 6 G; 0 T; 4 U; 0 Other;

Query Match 100.0%; Score 8; DB 2; Length 17;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 17 ATCGTCAG 10

RESULT 20

```

AAX71183/c
ID AAX71183 standard; RNA; 17 BP.
AC AAX71183;
XX
XX
DT 28-JUL-1999 (first entry)
XX
XX
DE Human KDR VEGF receptor hammerhead ribozyme substrate #195.
XX
XX
KW Vascular endothelial growth factor receptor; VEGF receptor; flt-1; flk-1;
KW KDR; hammerhead ribozyme; hairpin ribozyme; cleavage;
KW tumour angiogenesis; psoriasis; rheumatoid arthritis; ocular disease;
KW fms-like tyrosine kinase 1; kinase insert domain containing receptor;
KW foetal liver kinase 1; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN W09715662-A2.
XX
XX
PD 01-MAY-1997.
XX
XX
PF 25-OCT-1996; 96WO-US017480.
XX
XX
PR 26-OCT-1995; 95US-0005974P.
XX
PR 11-JAN-1996; 96US-00584040.
XX
XX
PA (RIBO-) RIBOZYME PHARM INC.
XX
PA (CHIR ) CHIRON CORP.
XX
XX
PI Pavco P, Mcswiggen J, Stinchcomb D, Escobedo J;
XX
XX
XX
DR WPI; 1997-259017/23.
XX
XX
PT Nucleic acid molecule modulating VEGF receptor(s) gene expression or mRNA
PT stability - useful for treating e.g. tumour angiogenesis, psoriasis,
PT rheumatoid arthritis, etc., in a human patient.
XX
XX
PS Claim 4; Page 103; 218pp; English.
XX
XX
CC The present invention describes nucleic acid molecules which modulate the
CC synthesis, expression and/or stability of a mRNA encoding 1 or more
CC receptors of vascular endothelial growth factor (VEGF). A patient
CC (preferably human) having a condition associated with the level of the
CC fms-like tyrosine kinase 1 (flt-1), kinase insert domain containing
CC receptor (KDR) and/or foetal liver kinase 1 (flk-1) (e.g. tumour
CC angiogenesis, ocular diseases, psoriasis and rheumatoid arthritis) can be
CC treated by administering the nucleic acid molecule or the expression
CC vector to the patient. AAX67275 to AAX75752 represent specific examples
CC of nucleic acid molecules from the present invention
XX
XX
SQ Sequence 17 BP; 5 A; 2 C; 5 G; 0 T; 5 U; 0 Other;

Query Match 100.0%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 8 ATCGTCAG 1

RESULT 21
AAX70610/c
ID AAA70610 standard; DNA; 17 BP.
XX
XX
AC AAX70610;
XX
XX
DT 15-SEP-2003 (revised)
DT 06-DEC-2000 (first entry)
XX
XX
DE Sindbis-like virus strain XJ-160 complete genome primer F1113.
XX
XX
KW Genome; Sindbis-like virus strain XJ-160; primer; RT-PCR; vaccine;
epidemic; Sindbis encephalitis; evolution; epidemiology; ss.
Sindbis-like virus; strain XJ-160.
CN1252444-A.
10-MAY-2000.
27-OCT-1998; 98CN-00120692.
27-OCT-1998; 98CN-00120692.
(VIRO-) INST VIROLOGY CHINESE ACAD PREVENTIVE ME.
Liang G, Li L, Zhou G;
WPI; 2000-443225/39.
Whole genome sequence of Sindbis virus strain and its cloning method.
Claim 3; Page 6; 17pp; Chinese.
Primers AAA70608-A70635 were used to RT-PCR amplify the complete genome
of the Sindbis-like virus strain XJ-160 (AAA70607). The genome was cloned
as 15 fragments using these PCR primers for inclusion into the plasmid
pGEM-T. The invention relates to the isolation and method of cloning the
complete genome for the Sindbis-like virus strain XJ-160 by a RT-PCR
process. The XJ-160 strain virus appears to be the optimal candidate for
a vaccine to prevent epidemics of Sindbis encephalitis. The sequence of
this strain's genome shows the difference between this viral strain and
other epidemic Sindbis virus strains at the molecular level and is useful
for understanding the source, evolution and molecular epidemiology of
Sindbis viruses. (Updated on 15-SEP-2003 to standardise OS field)
SQ Sequence 17 BP; 4 A; 6 C; 3 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 8; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 13 ATCGTCAG 6

RESULT 22
AAA25898
ID AAA25898 standard; DNA; 17 BP.
XX
XX
AC AAA25898;
XX
XX
DT 19-JUL-2000 (first entry)
XX
XX
DE Oestrogen receptor hammerhead ribozyme target sequence SEQ ID NO:2396.
XX
XX
KW Oestrogen receptor; c-raf; k-ras; bcl-2; ribozyme; cleavage;
KW hammerhead ribozyme; hairpin ribozyme; antisense oligonucleotide;
KW gene expression modification; cancer; phosphorothioate; endonuclease;
KW anticancer; breast cancer; endometrium cancer; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN W09954459-A2.
XX
XX
PD 28-OCT-1999.
XX
XX
PF 19-APR-1999; 99WO-US008547.
XX
PR 20-APR-1998; 98US-0082404P.
PR 23-JUN-1998; 98US-00103636.
XX
XX
PA (RIBO-) RIBOZYME PHARM INC.
XX
XX

```


Thompson JD, Beigelman L, Mcswiggen JA, Karpeisky A, Bellon L;
 Reynolds M, Zwick M, Jarvis T, Woolf T, Haerberli P;
 Matulic-Adamic J;
 WPI; 2000-013248/01.
 New nucleic acids that interact, and optionally cleave, target sequences,
 used to treat cancer.
 Claim 77; Page 94; 149pp; English.
 The present invention describes nucleic acids (A) that interact stably
 with a target sequence and contain at least one phosphoro(di)thioate
 link, having endonuclease activity. (A), and more generally any catalytic
 nucleic acid (A') that modulates expression of the oestrogen receptor
 gene, are used to treat cancer (particularly of breast or endometrium),
 in vivo or by transforming cells ex vivo and implanting treated cells, or
 for other conditions associated with levels of oestrogen receptor.
 Because of the high selectivity for targeted RNA, (A) can also be used to
 correlate inhibition of gene expression with alterations in phenotype,
 particularly for identification of therapeutic targets, and as research
 reagents (for RNA, in the same way that restriction endonucleases are
 used with DNA). The combination of modifications in (A) improves
 resistance to nucleases, binding affinity and/or activity. AAA23503 to
 AAA24747 represent oestrogen receptor hammerhead ribozyme sequences, and
 AAA24748 to AAA25992 represent their corresponding target sequences.
 AAA25993 to AAA26105 represent oestrogen receptor hairpin ribozyme
 sequences, and AAA26107 to AAA26218 represent their corresponding target
 sequences. AAA26219 to AAA26271 represent other ribozyme sequences and
 antisense oligonucleotides used in the exemplification of the present
 invention
 SQ Sequence 17 BP; 2 A; 5 C; 5 G; 5 T; 0 U; 0 Other;
 Query Match 100.0%; Score 8; DB 3; Length 17;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCGTCAG 8
 Db 4 ATCGTCAG 11
 RESULT 23
 AAA25897
 ID AAA25897 standard; DNA; 17 BP.
 AC AAA25897;
 AT 19-JUL-2000 (first entry)
 DE
 DE Oestrogen receptor hammerhead ribozyme target sequence SEQ ID NO:2395.
 KW Oestrogen receptor; c-raf; k-ras; bcl-2; ribozyme; cleavage;
 KW hammerhead ribozyme; hairpin ribozyme; antisense oligonucleotide;
 KW gene expression modification; cancer; phosphorothioate; endonuclease;
 KW anticancer; breast cancer; endometrium cancer; ss.
 OS Homo sapiens.
 XX
 XX WO9954459-A2.
 PN
 PD 28-OCT-1999.
 PF 19-APR-1999; 99WO-US008547.
 XX
 XX 20-APR-1998; 98US-0082404P.
 PR 23-JUN-1998; 98US-00103636.
 XX
 XX (RIBO-) RIBOZYME PHARM INC.
 PA
 PA Thompson JD, Beigelman L, Mcswiggen JA, Karpeisky A, Bellon L;
 PI Reynolds M, Zwick M, Jarvis T, Woolf T, Haerberli P;

Matulic-Adamic J;
 WPI; 2000-013248/01.
 New nucleic acids that interact, and optionally cleave, target sequences,
 used to treat cancer.
 Claim 77; Page 94; 149pp; English.
 The present invention describes nucleic acids (A) that interact stably
 with a target sequence and contain at least one phosphoro(di)thioate
 link, having endonuclease activity. (A), and more generally any catalytic
 nucleic acid (A') that modulates expression of the oestrogen receptor
 gene, are used to treat cancer (particularly of breast or endometrium),
 in vivo or by transforming cells ex vivo and implanting treated cells, or
 for other conditions associated with levels of oestrogen receptor.
 Because of the high selectivity for targeted RNA, (A) can also be used to
 correlate inhibition of gene expression with alterations in phenotype,
 particularly for identification of therapeutic targets, and as research
 reagents (for RNA, in the same way that restriction endonucleases are
 used with DNA). The combination of modifications in (A) improves
 resistance to nucleases, binding affinity and/or activity. AAA23503 to
 AAA24747 represent oestrogen receptor hammerhead ribozyme sequences, and
 AAA24748 to AAA25992 represent their corresponding target sequences.
 AAA25993 to AAA26105 represent oestrogen receptor hairpin ribozyme
 sequences, and AAA26107 to AAA26218 represent their corresponding target
 sequences. AAA26219 to AAA26271 represent other ribozyme sequences and
 antisense oligonucleotides used in the exemplification of the present
 invention
 SQ Sequence 17 BP; 5 A; 5 C; 3 G; 4 T; 0 U; 0 Other;
 Query Match 100.0%; Score 8; DB 3; Length 17;
 Best Local Similarity 100.0%; Pred. No. 4.6e+04;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCGTCAG 8
 Db 7 ATCGTCAG 14
 RESULT 24
 AEN08761
 ID AEN08761 standard; DNA; 17 BP.
 AC AEN08761;
 AT 29-MAY-2002 (first entry)
 DE
 DE Human GDMPL-1 17-mer scanning SEQ ID NO:5 sequence SEQ ID NO:8753.
 KW Human; genome-derived myosin-like protein 1; GDMPL-1; hGDMPL-1; heart;
 KW muscle; myosin; chromosome 22; gene therapy; vaccine; heart disease;
 KW skeletal muscle disorder; amplicon; screening; ss.
 OS Homo sapiens.
 XX
 XX WO200192524-A2.
 PN
 PD 06-DEC-2001.
 PF 25-MAY-2001; 2001WO-US016981.
 XX
 XX 26-MAY-2000; 2000US-0207456P.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 PR 30-JAN-2001; 2001WO-US000661.
 PR 30-JAN-2001; 2001WO-US000662.
 PR 30-JAN-2001; 2001WO-US000663.
 PR 30-JAN-2001; 2001WO-US000664.
 PR 30-JAN-2001; 2001WO-US000665.
 PR 30-JAN-2001; 2001WO-US000666.

30-JAN-2001; 2001WO-US000667.
30-JAN-2001; 2001WO-US000668.
30-JAN-2001; 2001WO-US000669.
30-JAN-2001; 2001WO-US000670.
05-FEB-2001; 2001US-0268660P.
(AEOM-) AEOMICA INC.
Gu Y, Ji Y, Penn SG, Hanzel DK, Rank DR, Chen W, Shannon ME;
WPI; 2002-179446/23.
New polypeptide, for raising antibodies that recognize hGDMPLP-1 proteins,
or as specific biomolecule capture probes for surface-enhanced laser
desorption ionization, comprises human myosin-like protein hGDMPLP-1.
Disclosure; SEQ ID NO 8753; 214pp; English.
The present invention describes a human genome-derived myosin-like
protein 1 (hGDMPLP-1). The protein and polynucleotide sequences of hGDMPLP-
1 can be used in gene therapy and vaccine production. The hGDMPLP-1
nucleic acids can be used as probes to detect, characterise and quantify
hGDMPLP-1 nucleic acids in samples, as amplification substrates, to
provide initial substrates for the recombinant engineering of hGDMPLP-1
protein variants having desired phenotypic improvements, and for
expressing the proteins. The hGDMPLP-1 proteins or polypeptides may be
used as immunogens to raise antibodies that specifically recognise hGDMPLP
-1 proteins, as standards in assays used to determine the concentration
and/or amount specifically of hGDMPLP proteins, as specific biomolecule
capture probes for surface-enhanced laser desorption ionisation, as
therapeutic supplement in patients having specific deficiency in hGDMPLP-1
production, and in vaccines or for replacement therapy. The
polynucleotide sequences encoding hGDMPLP-1 may be used for diagnosing a
disorder associated with the expression of hGDMPLP-1, in particular heart
and skeletal muscle disorders. hGDMPLP-1 is localised to chromosome 22.
The present sequence represents an oligomer used in the screening of the
hGDMPLP-1 sequence in the exemplification of the present invention. N.B.
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequence
Sequence 17 BP; 4 A; 6 C; 5 G; 2 T; 0 U; 0 Other;
Query Match 100.0%; Score 8; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 7 ATCGTCAG 14
RESULT 25
ABN08764
ID ABN08764 standard; DNA; 17 BP.
AC ABN08764;
XX 29-MAY-2002 (first entry)
DE Human GDMPLP-1 17-mer scanning SEQ ID NO:5 sequence SEQ ID NO:8756.
XX Human; genome-derived myosin-like protein 1; GDMPLP-1; heart;
KW muscle; myosin; chromosome 22; gene therapy; vaccine; heart disease;
KW skeletal muscle disorder; amplicon; screening; ss.
XX Homo sapiens.
XX WO200192524-A2.
XX 06-DEC-2001.
XX 25-MAY-2001; 2001WO-US016981.
26-MAY-2000; 2000US-0207456P.
21-SEP-2000; 2000US-0234687P.
27-SEP-2000; 2000US-0236353P.
04-OCT-2000; 2000GB-00024263.
30-JAN-2001; 2001WO-US000661.
30-JAN-2001; 2001WO-US000662.
30-JAN-2001; 2001WO-US000663.
30-JAN-2001; 2001WO-US000664.
30-JAN-2001; 2001WO-US000665.
30-JAN-2001; 2001WO-US000666.
30-JAN-2001; 2001WO-US000667.
30-JAN-2001; 2001WO-US000668.
30-JAN-2001; 2001WO-US000669.
30-JAN-2001; 2001WO-US000670.
05-FEB-2001; 2001US-0268660P.
(AEOM-) AEOMICA INC.
Gu Y, Ji Y, Penn SG, Hanzel DK, Rank DR, Chen W, Shannon ME;
WPI; 2002-179446/23.
New polypeptide, for raising antibodies that recognize hGDMPLP-1 proteins,
or as specific biomolecule capture probes for surface-enhanced laser
desorption ionization, comprises human myosin-like protein hGDMPLP-1.
Disclosure; SEQ ID NO 8756; 214pp; English.
The present invention describes a human genome-derived myosin-like
protein 1 (hGDMPLP-1). The protein and polynucleotide sequences of hGDMPLP-
1 can be used in gene therapy and vaccine production. The hGDMPLP-1
nucleic acids can be used as probes to detect, characterise and quantify
hGDMPLP-1 nucleic acids in samples, as amplification substrates, to
provide initial substrates for the recombinant engineering of hGDMPLP-1
protein variants having desired phenotypic improvements, and for
expressing the proteins. The hGDMPLP-1 proteins or polypeptides may be
used as immunogens to raise antibodies that specifically recognise hGDMPLP
-1 proteins, as standards in assays used to determine the concentration
and/or amount specifically of hGDMPLP proteins, as specific biomolecule
capture probes for surface-enhanced laser desorption ionisation, as
therapeutic supplement in patients having specific deficiency in hGDMPLP-1
production, and in vaccines or for replacement therapy. The
polynucleotide sequences encoding hGDMPLP-1 may be used for diagnosing a
disorder associated with the expression of hGDMPLP-1, in particular heart
and skeletal muscle disorders. hGDMPLP-1 is localised to chromosome 22.
The present sequence represents an oligomer used in the screening of the
hGDMPLP-1 sequence in the exemplification of the present invention. N.B.
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequence
Sequence 17 BP; 4 A; 6 C; 5 G; 2 T; 0 U; 0 Other;
Query Match 100.0%; Score 8; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 7 ATCGTCAG 14
Query Match 100.0%; Score 8; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.6e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 4 ATCGTCAG 11
Search completed: March 11, 2005, 04:19:48
Job time : 15.7034 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:06:25 ; Search time 100.882 Seconds
(without alignments)
8645.645 Million cell updates/sec

Title: US-09-674-277-20
Perfect score: 18
Sequence: 1 accgcatcgctcagttgsg 18

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	6	BD205230 Nucleotid
2	18	100.0	18	6	AX011316 Sequence
3	18	100.0	31	6	BD205222 Nucleotid
4	18	100.0	31	6	BD205223 Nucleotid
5	18	100.0	31	6	AX011308 Sequence
6	18	100.0	31	6	AX011309 Sequence
7	18	100.0	1489	6	BD205211 Nucleotid
8	18	100.0	1489	6	AX011297 Sequence
9	18	100.0	92077	1	AF074613 Escherich
10	18	100.0	92077	6	AX191727 Sequence
11	18	100.0	92721	1	AB011549 Escherich
12	18	100.0	92721	6	AX191725 Sequence
13	17	94.4	1517	1	PSIS801
14	17	94.4	40110	1	AY603980
15	16.4	91.1	890	1	PSEPZH1
16	16.4	91.1	966	8	BT008719 Arabidops
17	16.4	91.1	1137	6	AX506723 Sequence
18	16.4	91.1	1177	8	AY217003 Saccharom
19	16.4	91.1	1301	8	AY087588 Arabidops

20	16.4	91.1	1307	8	AF370481
c 21	16.4	91.1	1710	8	SCYBRI45W
22	16.4	91.1	198354	8	ATYB22
c 23	16.4	91.1	198750	8	ATCHRIV85
24	16.4	91.1	313846	1	AE017234 Mycobacte
c 25	16.4	91.1	349746	1	BX572605 Rhodopsea
26	16.4	91.1	349761	1	BX842646 Bdellovib
27	16	88.9	16	6	BD205229 Nucleotid
28	16	88.9	16	6	AX011315 Sequence
29	15.4	85.6	594	8	BZAIT5
c 30	15.4	85.6	1356	14	AF320813 Hepatit18
c 31	15.4	85.6	1632	6	A44905 Sequence 1
c 32	15.4	85.6	1632	6	I49891 Sequence 1
c 33	15.4	85.6	1632	6	I93610 Sequence 1
c 34	15.4	85.6	2207	8	AF332874 Oryza sat
c 35	15.4	85.6	2307	8	AK070452 Oryza sat
c 36	15.4	85.6	10937	1	AE012029 Xanthomon
c 37	15.4	85.6	13897	1	ABCLSDA
c 38	15.4	85.6	20889	1	AE008822 Salmonell
c 39	15.4	85.6	22934	6	CQ363825 Sequence
c 40	15.4	85.6	25860	6	AX933596 Sequence
41	15.4	85.6	39554	1	AF359557 Pseudomon
c 42	15.4	85.6	110086	2	AP003831 Oryza sat
c 43	15.4	85.6	110000	1	AE000516 05
44	15.4	85.6	110000	1	AE000516_06
c 45	15.4	85.6	110000	1	AE017283_12

ALIGNMENTS

RESULT 1
BD205230

LOCUS

DEFINITION

BD205230 18 bp DNA linear PAT 17-JUL-2003
Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD205230 1 (bases 1 to 18)
Frechon,D.T.M., Laure,F.C. and Thierry,D.
Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)

Patent: JP 2002512813-A 20 08-MAY-2002;
BIORAD PASTEUR
OS Unidentified
PN JP 2002512813-A/20
PD 08-MAY-2002
PF 27-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDE LAURE, PI

PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC

Strandedness: Single;
CC Topology: Linear;

CC Nucleotide sequence for detecting enterohemorrhagic CC

CC Escherichia coli

CC (EHEC).

FT source

FT key

FT location/Qualifiers

FT 1..18

FT /organism='Unidentified'.

FEATURES

source

1..18

/organism='unidentified'

/mol_type='genomic DNA'

/db_xref='taxon:32644'

ORIGIN

Query Match 100.0% Score 18; DB 6; Length 18;

Best Local Similarity 100.0% Pred. No. 2.7e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 1 ACGGCATCGTCAGTTGCG 18

RESULT 2
AX011316
LOCUS AX011316 18 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 20 from Patent WO9955908.
ACCESSION AX011316
VERSION AX011316.1 GI:9997866
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1
REFERENCE Thierry, D., Frechon, D.T. and Laure, F.C.
AUTHORS Nucleotide sequences for detecting enterohemorrhagic escherichia
TITLE coli (ehc)
JOURNAL Patent: WO 9955908-A 20 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
FRANCOISE CLAUDINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES Location/Qualifiers
source 1..18
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 1 ACGGCATCGTCAGTTGCG 18

RESULT 3
BD205222
LOCUS BD205222 31 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION BD205222
VERSION BD205222.1 GI:33014992
KEYWORDS JP 2002512813-A/12.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 31)
AUTHORS Frechon,D.T.M., Laure,F.C. and Thierry,D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 12 08-MAY-2002;
BIORAD PASTEUR
COMMENT OS Unidentified
PN JP 2002512813-A/12
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FT source 1..31
Location/Qualifiers

FEATURES
source Location/Qualifiers
1..31
/organism="Unidentified".
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 5 ACGGCATCGTCAGTTGCG 22

RESULT 4
BD205223
LOCUS BD205223 31 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION BD205223
VERSION BD205223.1 GI:33014993
KEYWORDS JP 2002512813-A/13.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 31)
AUTHORS Frechon,D.T.M., Laure,F.C. and Thierry,D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 13 08-MAY-2002;
BIORAD PASTEUR
COMMENT OS Unidentified
PN JP 2002512813-A/13
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FT source 1..31
Location/Qualifiers

FEATURES
source Location/Qualifiers
1..31
/organism="Unidentified".
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 9 ACGGCATCGTCAGTTGCG 26

RESULT 5
AX011308
LOCUS AX011308 31 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 12 from Patent WO9955908.
ACCESSION AX011308
VERSION AX011308.1 GI:9997858
KEYWORDS

SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Thierry, D., Frechon, D.T. and Laure, F.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)

JOURNAL Patent: WO 9955908-A 12 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES
source Location/Qualifiers
1..31
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 5 ACGGCATCGTCAGTTGCG 22
|||||

RESULT 6
AX011309
LOCUS AX011309 31 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 13 from Patent WO9955908.
ACCESSION AX011309
VERSION AX011309.1 GI:9997859
KEYWORDS
SOURCE
ORGANISM Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Thierry, D., Frechon, D.T. and Laure, F.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)

JOURNAL Patent: WO 9955908-A 13 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES
source Location/Qualifiers
1..31
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 9 ACGGCATCGTCAGTTGCG 26
|||||

RESULT 7
BD205211
LOCUS BD205211 1489 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).

ACCESSION BD205211
VERSION BD205211.1 GI:33014981
KEYWORDS JP 2002512813-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1489)

AUTHORS Frechon, D.T.M., Laure, F.C. and Thierry, D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 1 08-MAY-2002;
BIORAD PASTEUR

COMMENT OS Unidentified
PN JP 2002512813-A/1
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PI 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDE LAURE, PI DOMINIQUE THIERRY

PC C12N9/08, C07K14/245, C12N1/21, C12N15/09, C12Q1/68, C12N15/00 CC
Strandedness: Double;
CC Topology: linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC Escherichia coli (EHEC).

EH Key Location/Qualifiers
FT source 1..1489
/organism="Unidentified".

FEATURES
source Location/Qualifiers
1..1489
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 1489;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 395 ACGGCATCGTCAGTTGCG 412
|||||

RESULT 8
AX011297
LOCUS AX011297 1489 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent WO9955908.
ACCESSION AX011297
VERSION AX011297.1 GI:9997847
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Thierry, D., Frechon, D.T. and Laure, F.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)

JOURNAL Patent: WO 9955908-A 1 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES
source Location/Qualifiers
1..1489
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 1489;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 395 ACGGCATCGTCAGTTGCG 412
|||||

RESULT 9
AF074613

/gene="L7007"
/note="91 pct identical (0 gaps) to 24 residues of an
approx. 24 aa protein uORF P30REFFIC, accession M6167,
translationally coupled to replication initiation protein"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70075.1"
/db_xref="GI:3822121"
/translation="MLGKVQDFFLCSLLLRIVSAGWCD"
3032..3035
/note="100 pct identical (0 gaps) to RBS at 641..644 locus
ECNR1REP accession X02302"
3040..3897
/gene="rep2"
/note="synonym: L7008"
3040..3897
/gene="rep2"
/note="96 pct identical amino acid sequence and equal
length to REP2_ECOLI SM: P03066"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/translation="MTDLOQTYVROKVPNPVTPREGACTLKECEKMEKAVQFTSR
FDPAIHVAHRSKGLRRMPVLRRAIDALLQGLCHFYDPLANRVQCSITLAIKCG
LATESAAGKUSITRALTPLSGLDITYTEYDPLIGCYIPTDITFTTALFAALDV
SEDAVAARSRWENRKLQGLDITLGMDELIAKARFVRFRFSYOTELKSRGK
RARARDAGRQDITLVKRLQTLTSEGRFTANREAVKREVRVKERMILSRNRN
YSRLATASP"
3847..4095
/standard_name="CIS"
/note="88 pct identical to locus ECCIS accession X12587,
required for cis-activation of oriR by the replication
initiation protein"
4072..4080
/note="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1REP at (1682..1690) accession X02302"
4084..4232
/standard_name="oriR"
/note="89 pct identical to oriR (1094..1242); minimum
segment for replication of E. coli IncFII plasmid NR1
ECREPA1_X12776"
/direction=right
4189..4197
/note="predicted sigma 70 promoter; score of 56%"
4259..4453
/gene="L7009"
4259..4453
/gene="L7009"
/note="95 pct identical to (0 gaps) 64 residues of a 128
aa protein REPA4 locus ECRSCL accession V00351"
/codon_start=1

Query Match 100.0%; Score 18; DB 1; Length 92077;
Best Local Similarity 100.0%; Pred. NO. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 7231 ACGGCATCGTCAGTTGCG 7248

RESULT 10
AX191727
LOCUS AX191727 92077 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 9 from Patent WO0149775.
ACCESSION AX191727
VERSION AX191727.1 GI:15209896
KEYWORDS Escherichia coli
SOURCE Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1
Iversen,P.L.
Antisense antibacterial cell division composition and method
Patent: WO 0149775-A 9 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source
Location/Qualifiers
1..92077
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 18; DB 6; Length 92077;
Best Local Similarity 100.0%; Pred. NO. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 7231 ACGGCATCGTCAGTTGCG 7248
RESULT 11
AB011549
LOCUS 92721 bp DNA circular BCT 27-APR-1999
DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION AB011549
VERSION AB011549.2 GI:4589740
KEYWORDS ToxR-regulated lipoprotein; tagA.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (sites)
Makino,K., Ishii,K., Yasunaga,T., Hattori,M., Yokoyama,K.,
Yutsudo,H.C., Kubota,Y., Yamaichi,Y., Iida,T., Yamamoto,K.,
Honda,T., Han,C.G., Ohtsubo,E., Kasamatsu,M., Hayashi,T., Kuhara,S.
and Shinagawa,H.
Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
outbreak
DNA Res. 5 (1), 1-9 (1998)
98290540
9628576
2 (bases 1 to 92721)
Direct Submission
Makino,K.
Submitted (24-FEB-1998) Kozo Makino, Research Institute for
Microbial Diseases, Osaka University, Molecular Microbiology,
Yamadaoka, 3-1, Suita, Osaka 562, Japan
{E-mail:makino@bks01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
Fax:81-6-879-8320}
On Apr 20, 1999 this sequence version replaced gi:3336997.
COMMENT
Location/Qualifiers
1..92721
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:562"
/plasmid="pO157"
/note="RIMD 0509952 is a strain of enterohemorrhagic E.
coli, EHEC O157:H7"
join(92527..92721,1..2502)
/gene="tagA"
join(92527..92721,1..2502)
/gene="tagA"
/codon_start=1
/transl_table=11
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4666293"

```
/translation="MNTKKNRWRTPMKLYLSCTILAPLIGVFSATAADNNNAIYF
NTSQPINDQGSAAEVKKAQSQILPAHPKREGSQHLSLRSLILVLRPVKADKTP
VQVEARDNNKILTLTLVPSLPDTIYHLGDVPEGIDFTPHNGKKIINTVAEVN
KLSADSGSSIHSLHTNALVEIHTANGRWVRDIYLPQGPDEGMVRFVSSAGSYTV
FYGRKVTLSVGNLTLPKYNGQWFRSGELENRNRIYAQHIMSAELPAHWIVPGLNLV
IKQGNLSGRNLDIKIGAPGELLHTIDIGMLTTPRDRFDPAKDEAHREVFOTIPVSR
MIYNNVAPHLKVEMLPTGBELLTDMDPGNGGWHSCIMRQRIKGLSHGIDNANYGLN
STAGLGENSHPIYVAQAAHNSRGTANGIQVHGGSGGGIIVTLDSTLGNFSEVGH
NYGLHYVDGFGKSVHRSANNSTWGDGKKRFIPNFYPSQTNSEKSLCNLCQSPFF
DGHKFGDAMAGSPSAAANRFMTYIPNSAI IQRFENKAVFDSRSSTGFSKWNADT
QMEPEYHTIDRAEQITASVNELSEKMAELMAEYAVVKVMHNGWNRNIYIPTASA
DNKGSILTINHEAGYSYLPINGDEKVVSGYKKSFFVSDQGFWKERDVVDVTRAKPE
QGFVPVTLVGYDPEGLTSSYIYPAMYGAQGYTYSDDSQNLSDNCQLOVDTKEGOL
RFLANHRNNTVMNRHINVPTESTQTAIIVCNKILDTKSLTAPAREGLTYVNGQ
ALPAKENEGCIVNSGKRYCLPVGQRSGSLPDWIVGQEVYVDSGAKAKVLLUSDWDN
LSYNRIGEFVGNVPADMKKVKAWNGQYLDSPKPSRMRVYVY"
2589. .3464
/gene="etpC"
2589. .3464
CDS
/codon_start=1
/gene="etpC"
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="GI:3336993"
/translation="MLFPLSFRDGRGLFIKDIIVLKMLTNRLILCVILLIAGYQIVSVI
HFWLTOAASVPLGSLYSAPETAVTGQTEERFVFTLFGRASPLSSEGRAQETWPSLS
DILLSGDLVGRILYSVAEHSVAIFAHNRQFSLSGVSKFSPSYDATISAIFSDHIV
INQGVTSVLPDRYDTEKKNAYDNNLTVGDVITQDNFRVSEFDMISFSAVTVNNT
LSGVRLIPGKHSILFYNAGLHDNDLAVSVNGSELDRTRQAQIMKQLPKELKEITVE
RDQLYDAFTAIVEN"
3675. .5432
/gene="etpD"
3675. .5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="GI:3337000"
/translation="MLNEEQVYQFFLSVDVYFVAVVDMHNGILKVVRSKDAKTSVAP
VASVSPCTGDEVTVRVVPVSVNVAARDLAPLLROLNDNAGAGSVVHYEPSNVLLMTGR
AANVKRLMEIVERVDKVGNSVATVPLTYASATDVABLTELKETDKTALPAWMTAK
LVADERTNSVLSGEPISQQRITSIKQLDQEDVQGNKVIYLYKAKADLVEVLTG
ISSIENDSKKSPSTEALRKGVITIKSHQENALILTGAPDVIRLENVVISQLDTRRFQ
VLVEAIIAETQDADNLGIQWVNKHGAVQFTSGLPITTMVQTRQNEILSDQNSA
LSMFMGIAAGFYQGNWAMLLTALSTSSKNIDILATPSIVTLDNMEATFNVOQEVPLSG
SOTTSQGNIFNTVERKTVGILRVKPOINEGDSVLLBIEQVSGVADTAVATVTDLGA
TFNTRVTNAMLVNGETVVVGGLLDKSIRGSESKVPLLGDIPLVGLHLFRAKSBQTA
RNLMFRTIIRDRGDFRHSABEKYSQFNQEQVQSRGKETTALTLINEEQRLRLSPDQD
DTAFPRKVAALFAFYAQM"
5432. .6937
/gene="etpE"
5432. .6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="GI:3337001"
/translation="MSRVQNVSVSRPLLPSPFSRTORILLRBEQGNRVFCMBDTPA
SALLEVRVAEGPLNVTIVSABAEKQLVSSYQRDSDEARQMMAEIAGNEMDFYTVAGE
LPDRDLDLNDADAPIRLINAMLTAEIKKASDILHETIERHQLQFRFDIGVLRREIL
LRHLNLSLISRIKVMARLDIAEKRPQDGMWRILRGRAVDVVRVSTPLSNHGERIV
LRLLDKNSVLDLAALGMSQONQHIDALIRRHGIIILVGTGSGSKSTTLVLAALSLL
NPRDNTWTVEDPVEYELDGI SQTVNPKVMTFARSILRLQDPDQVSLVGEIRDGE
TAQIAVQASLTHLVLSTLNTNSAGALSRLQDMGIPPELSTSLLAVALQRLVRLTC
PRCQPCQVSTELANDMDIIPETTINGQKQVLAGESTWEEVVRVRESLQEEE"
6939. .8162
/gene="etpF"
6939. .8162
/gene="etpF"
CDS
/gene="etpF"
```

```
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="BAA31761.1"
/translation="MALFYQASDIHGRKRSGLSEADSAHARQQLREQALIPVRLDE
KOVHKKHSILSRFRPGGSAELALTRQLATLVAASPLEBEALDALLRQSEKPRQ
RNLLAAVRTKLEHSLAAAMGMPGCTFERLYCAMVAAGETSRLDVLISRLADYTEQ
RQIMRNLLOALLYPCVLTVAVGVTAIILLTAVPVKVQBFIMHKQTLPLSTRVLWGA
AEVQTSWGWALLAALGALIGRMILHQBSRLAFHLLRLRPVVGRIISRGLTARYA
RTLSILNAPAVLLOQWHISGDVLSNDWARHQIATAAELVREGVSLHQALETSLPFP
MWRHMTASGENSGELDSMLERAADNQDRFSTQMQLALGLFEPFLVVGMAGVVILFVL
AIIQPLLIQNNMMNM"
8193. .8627
/gene="etpG"
8193. .8627
/gene="etpG"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="BAA31762.1"
/translation="MRKQHQGFLLIEMVIVILGVLASLVVPNLMGNKDKADROKV
MSDLVALESTLDMYRLDNNRYPTTEQGLRALVSKPTVQPEPNRYQDGIYRRLPQDDPW
GGDYQLLNPCQYSDIDIFSPGPDGVPTEDDIGNWTIGNAQP"
8624. .9178
/gene="etpH"
8624. .9178
/gene="etpH"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="GI:3337003"
/translation="MSQRGFTILEMMILLIMCITAGLVLMSPFDSNQNHLOQOREL
QOQDYALDRSQDGLLMGIQVRFDGWKFKVLQGRTAESPPTLAEGGDIWQGVFWQWV
QPRRAAMGGKLPPDVRLQLQYLRGLQWMSEHDDGAEPIILLPGGEVTPFRLLLFRQVG
EBAVVGLOVDENGLMTLFEGEVSL"
9193. .9540
/gene="etpI"
9193. .9540
/gene="etpI"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="BAA31764.1"
/translation="MTLLEVIIVAVFALGAMLMQASQOAAAGIGRMEBKVLAWILA
DNQMVQLQLEKTPWENGGEKTIISFAGTEWYLRWGGPDSVDPQPSRSEVEVRRKTEET
SALVLSRSSVVRE"
9537. .10136
/gene="etpJ"
9537. .10136
/gene="etpJ"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="BAA31765.1"
/translation="MSQQRVKGFTLLEMLALAVFAALISAFQVLQGGIRAHLSRD
KVORLABLQGISQMRDRLTQMLPRHSRNEGGLLAAAPHLLKSLDDWGISFTNRSWLN
AGMLPRDELQWGYRLRQKLERLSYFPHVDHPSGVSPDVRLMDGVHAFRLRFFVNGD
WQARWDSGTGLPQAVEVTLVMDDFAEPLRPLFLVSKETA"
10133. .11110
/gene="etpK"
10133. .11110
/gene="etpK"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/db_xref="BAA31766.1"
/translation="MKUREQGVALLVLLILSLMVTIAAVERNGRTFLRTVAQLQDQ
```


LOAKWDGYTAETIAKQILQSRQOESPRKTHLAONWAQSEROFETRGGDVRGQIVDAQA
CFNLNAINGVVDLTSIPYARIQOLLINQVELLQARQVTAALRDWIDDDXRVRG
GAEBVYMGPPILAAQPMQDVSLELRIGIDARLYRKLLPYCVLPDTSLSNVVN
TLDSQAPLLAALFLTPDLSFVTELLQRRPTGWESVAAFLDAPLDIDTSAAMPV
LAVSSNYFLVRLHVRSGEHLFSQOTLMQWREERFRIIQRYGLTMREVP"
11473. .12321

gene
Query Match 100.0%; Score 18; DB 1; Length 92721;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 76746 ACGGCATCGTCAGTTGCG 76763

RESULT 12
AX191725 92721 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 7 from Patent WO0149775.
ACCESSION AX191725
VERSION AX191725.1 GI:15209894

KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Iversen, P. L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)

FEATURES
source
1. .92721
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 92721;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 76746 ACGGCATCGTCAGTTGCG 76763

RESULT 13
PSIS801 1517 bp DNA linear BCT 07-JUL-2002
LOCUS
DEFINITION P.syringae DNA for IS801 insertion sequence.
ACCESSION X57269
VERSION X57269.1 GI:45830

KEYWORDS insertion element IS801.
SOURCE Pseudomonas syringae
ORGANISM Pseudomonas syringae
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE 1
AUTHORS Romantschuk, M., Richter, G. Y., Mukhopadhyay, P. and Mills, D.
TITLE IS801, an insertion sequence element isolated from Pseudomonas
syringae pathovar phaseolicola
JOURNAL Mol. Microbiol. 5 (3), 617-622 (1991)

MEDLINE 91260445
PUBMED 1646375
REFERENCE 2 (bases 1 to 1517)

AUTHORS Mills, D.
TITLE Direct Submission

JOURNAL Submitted (23-JAN-1991) D. Mills, Oregon State University, Dept of
Botany and Plant Pathology, Corvallis OR 97331-2902, U S A
FEATURES
source
1. .1517
Location/Qualifiers

/organism="Pseudomonas syringae"
/mol_type="genomic DNA"
/strain="pathovar phaseolicola, strain LR781"
/db_xref="taxon:317"

misc_feature
1. .5
/note="insertion target sequence duplication"

repeat_region
6. .1512
/evidence=experimental
/insertion_seq="IS801"

CDS
101. .1333
/note="unnamed protein product; orf1"
/codon_start=1
/transl_table=11
/protein_id="CAA40540.1"

/db_xref="GI:45831"
/db_xref="GOA:P24607"
/db_xref="UniProt/Swiss-Prot:P24607"
/translation="MKPAYPLQLQMSCHPHKYLCLNTCHCRACPCGKKATQWITVON
VEVTKMLACGTSILGVKHYTCGNHSCPHKYLCLNTCHCRACPCGKKATQWITVON
NRLPDCQWHLVFTLPDLPLFFYNRLWLDALFRLADNLIYAAKRGRLVGFAL
HTYGRNLNHPHVLSTAGLDDEQGVKNLSFKHEALRRKRWMLVRLYLGLQPLSQL
TMPPLAHILCESDMRLILAAAGQWHIHLSSKTKNGRTVNYLGRYKKPPISSGR
LAHYTGATILRFTYLDHRTQAYQOETLSQADMLFRVVQHIPEKHFRMIRYFGFLANRV
CGOYLPKVYBALKMATPGPTPKLYFAPMAKAFNLVDPFRCLCGARWVYTAISGLTV
QGLNLNQAIAQMRVYKP"

RBS
387. .391
/note="putative"
complement (461. .1060)
/note="unnamed protein product; orf2"
/codon_start=1
/transl_table=11
/protein_id="CAA40541.1"

/db_xref="GI:45832"
/db_xref="UniProt/TrEMBL:Q99178"
/translation="MLFRDLVHHPKHVGLAQRELILGLCAVIOVGEPQPGVGMVRO
TTARDGRFFQVAAQVIDGLSAVPELLRQVDVPLVLTASQVOTAPVAFTEDEVGQRRHR
QLKRLSQVIAHQHPAPPGLEVERQIIFDALLIOAARGHRQVDVGMPEVPEPAIGV
QRPNPDATAPLGGVDVIGRQPKRIEQPAVVEEQGP"

misc_feature
1513. .1517
/note="insertion target sequence"
/evidence=experimental

ORIGIN

Query Match 94.4%; Score 17; DB 1; Length 1517;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 763 CGGCATCGTCAGTTGCG 747

RESULT 14

LOCUS

AY603980

40110 bp DNA

circular BCT 20-JUL-2004

Pseudomonas syringae pv. maculicola strain ES4326 plasmid

PPMA4326B, complete sequence.

ACCESSION AY603980

VERSION AY603980.1

GI:47525154

KEYWORDS

Pseudomonas syringae pv. maculicola

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 40110)

AUTHORS Stavrinides, J. and Guttman, D. S.

TITLE Nucleotide Sequence and Evolution of the Five-Plasmid Complement of

the Phytopathogen Pseudomonas syringae pv. maculicola ES4326

JOURNAL J. Bacteriol. 186 (15), 5101-5115 (2004)

PUBMED 15262947

REFERENCE 2 (bases 1 to 40110)

AUTHORS Stavrinides, J.

TITLE Direct Submission


```
ESISKMLACGTSILGVKHYTCANBCHPHVKYLCNTCHRCACPCSGCKKATDQWIAVQNN
RLDPCQHVLVFTLPDILWLPFFYFNWLLDALPRLAADNLIYAARKRGLRVGIFGALH
TYGRLNWHRHVHLSVTAGDDEGVWKNLSFHEEARLRRMMWLVRDYLILGQPSQUT
MPPQAHILSDESRRDITAGGQWHIHSKTPGKTKVNYLGRYKLPFGSGSL
AHTYCGATLSFTYDHRITQYQETLSQADMLRVQVHIPEKHFMRIRYFGLANRVC
GOYLPKYVEALKMATPGVPKLYFAQWAKAFNLVDPFRCVLCGARMVVTAAISGLTVQ
GLVNAQAIQMRVVKP"
complement (8608) .9330)
/locus_tag="PMA4326808"
CDS
complement (8608) .9330)
/locus_tag="PMA4326808"
/notes="similar to Pseudomonas syringae pv. maculicola
ES4326 type III effector HopPmaA encoded by GenBank
Accession Number AF458040"
/codon_start=1
/transl_table=11
/product="candidate type III effector"
/protein_id="AAT35175.1"
/db_xref="GI:47525162"
/translacion="MSPAQIIRTSHPSPSTGTSSAENSHAQSPQOVLTRAFVSG
ELNARFORTSTASQDFTSLGLTQRLSEKRTFAPDLAEALANQALHQAAGDQGGHNL
GRDEQQLTGMIDRCKSQLAHTPASDYSYLLAQVCNLRKARLHQSIQMTGEAHAK
VRGVPDILAIQDPLDAEKPGVMPGVYKFSFICMAKARTAELENLRNASNEVAL
LLHPHADTILEQGSRLQIVGF"
misc_feature
9898..10196
/notes="element 622; E622; MFB03"
gene
10161..10766
/locus_tag="PMA4326809"
CDS
10161..10766
/locus_tag="PMA4326809"
/notes="similar to Salmonella typhimurium LT2 plasmid pSLT
putative resolvase encoded by GenBank Accession Number
AE006471"
/codon_start=1
/transl_table=11
Query Match 94.4%; Score 17; DB 1; Length 40110;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 7632 CGGCATCGTCAGTTGCG 7648

PSEPHZ1
LOCUS 890 bp DNA linear BCT 24-JUN-1996
DEFINITION Pseudomonas aureofaciens autoinducer synthase (phzI) gene, complete
cda.
ACCESSION L33724.1 GI:499662
VERSION L33724.1
KEYWORDS autoinducer synthase; phzI gene.
SOURCE Pseudomonas chlororaphis
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 890)
AUTHORS Wood,D.W. and Piersen,L.S. III.
TITLE The phzI gene of Pseudomonas aureofaciens 30-84 is responsible for
the production of a diffusible signal required for phenazine
antibiotic production
JOURNAL Gene 168 (1), 49-53 (1996)
MEDLINE 96186954
PUBMED 8626064
COMMENT Original source text: Pseudomonas aureofaciens (strain 30-84) DNA.
FEATURES
source
Location/Qualifiers
1..890
/organism="Pseudomonas chlororaphis"
/mol_type="genomic DNA"
/strain="30-84"
/db_xref="taxon:333"
133..796
gene
```

```
/gene="phzI"
133..140
/genes="phzI"
/notes="putative"
164..169
/genes="phzI"
/notes="putative"
191..196
/genes="phzI"
/notes="putative"
206..796
/genes="phzI"
/notes="putative"
/codon_start=1
/transl_table=11
/product="autoinducer synthase"
/protein_id="AAC41535.1"
/db_xref="GI:499663"
/translacion="MMMEHTLNMDSDELKMLGRPHRQFVEKGLRMLPAHPQOAGC
ENQOYDEHARYLLAFNEDRAIVGCARLIPTTFNLLGCVFGHTCAGAPKPKHAIWEM
TRETTEPQAMPFLFWRSKLTASLAGADAIVGVNSTMYRYKINGVHYERKLGVPVTVH
ONEKILAIKLSAHRHRSVAVPSAFMSDTLLRETA"
ORIGIN
Query Match 91.1%; Score 16.4; DB 1; Length 890;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 551 ACGGCATCGTCAGTTGCG 534

RESULT 16
LOCUS BT008719 966 bp mRNA linear PLN 22-MAY-2003
DEFINITION Arabidopsis thaliana At4g36530 gene, complete cds.
ACCESSION BT008719
VERSION BT008719.1 GI:30984537
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
REFERENCE 1 (bases 1 to 966)
AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T.,
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 966)
AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T.,
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2003) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
```

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/salk) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
1..966
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U15142"
/ecotype="Columbia"
/note="This clone is in pUNI 51"

CDS

1..966
/note="unknown protein"
/codon_start=1
/product="At4g36530"
/protein_id="AAP42732.1"
/db_xref="GI:30984538"
/translation="MDTSSASQSVQSERLAFKPEGYNFWRGKHIVYVQEGSGP
LVTHFGCASVFWRYNIPELAKKYVYALDILGFGWSDKALIEYDAMVTDQVIDFM
KEVVEPAVVVGLSGVLTALSVAGLPEQVTVGALLNSAGQFAAESRKEEAEDEV
TKETVKPLKEIFORVILGELFMQAKPSRIETSLKSVYIDSTNVDDVYLVESISKPAVD
PNAGEVYRLMTRFLTNQSRITLDSVLSKTCPLLLVWGLDPMWGPAPAKKIKAFYS
NSSLVHLQAGCHPDEPVEAVNKALLDWLSINIAKSPASPIVLET"

ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 966;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

Db 361 ACGGCATGTCAGTTGCG 378

RESULT 17
AX506723
LOCUS AX506723 1137 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1418 from Patent WO0216655.
ACCESSION AX506723
VERSION AX506723.1 GI:23387960

KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana

REFERENCE 1
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use
JOURNAL Patent: WO 0216655-A 1418 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG (CH)

FEATURES
source
1..1137
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 91.1%; Score 16.4; DB 6; Length 1137;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

Db 532 ACGGCATGTCAGTTGCG 549

RESULT 18
AY217003/c
LOCUS AY217003 1177 bp DNA linear PLN 01-FEB-2004
DEFINITION Saccharomyces pastorianus alcohol dehydrogenase 5 (ADH5) gene, complete cds.

ACCESSION AY217003
VERSION AY217003.1 GI:37787880
KEYWORDS Saccharomyces pastorianus
SOURCE Saccharomyces pastorianus
ORGANISM Saccharomyces pastorianus

REFERENCE 1 (bases 1 to 1177)
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
AUTHORS Thomson, J.M.
TITLE Interpretive Proteomics: Experimental paleogenetics as a tool to analyze function and discover pathways in yeast
JOURNAL Thesis (2002) University of Florida
REFERENCE 2 (bases 1 to 1177)
Thomson, J.M. and Benner, S.A.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2003) Anatomy & Cell Biology, University of Florida College of Medicine, 1600 SW Archer Rd., Rm. B1-3, Gainesville, FL 32610, USA

FEATURES
source
1..1177
/organism="Saccharomyces pastorianus"
/mol_type="genomic DNA"
/db_xref="taxon:27292"
<65..>11120
/gene="ADH5"
<65..>11120
/genes="ADH5"
/product="alcohol dehydrogenase 5"
65..1120
/gene="ADH5"
/EC number="1.1.1.1"
/note="SpADH5"
/codon_start=1
/product="alcohol dehydrogenase 5"
/protein_id="AAP51053.1"
/db_xref="GI:37787881"
/translation="MPSQVLPBKQKAIYFETDGKLEYKDVTPPEKPNELVHVKYS
GVCHSDLHAMHGWPFQFOLKPLIGGHEGAGVWVKGNSVKGWGVDFAGIKWLNGTCM
SCVEYEVNESQCPYLDGTGTHDGTFOEVATADAVQAAHI PPNVNLAEVAPILCAGI
TVYKALKRANVLPQWVTISGACGGLSLAIOVALAGYRVIIGIDGNAKRLFEQLG
GEIFDFTBEKDIVGAILIKATIGSGHGVINVSSEALEASTRYCRPNGTVLVGMFA
HAYCNSDVFNVKVSISVSCVGNRADTREALDFFARGLIKSPHLAGLSDVPEIFA
KMEKGEIVGRYVYVETSK"

ORIGIN
Query Match 91.1%; Score 16.4; DB 8; Length 1177;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

Db 477 ACGGCATCGTCAGTTGCG 460

RESULT 19
AY087588
LOCUS AY087588 1301 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 36845 mRNA, complete sequence.

ACCESSION AY087588
VERSION AY087588.1 GI:21406326
KEYWORDS FLI_CDNA.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1301)
Haas, B.J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.

TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22088475
PUBMED 12093376

REFERENCE 2 (bases 1 to 1301)
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.

TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1301)
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and Feldmann, K.

TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. GenSet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES
source
Location/Qualifiers
1..1301
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clones="36845"
42..1178
/codon_start=1
/product="unknown"
/protein_id="AAM5130.1"
/db_xref="GI:21593181"
/translation="MSASCALTPSVRTELFSSSSKRSIFPARICRPNKCEISRDF AIRGGIVASGVNDTSSASQSVQSERLAFKPEGYNFWWRGKHIVYVQVGGESPL VLTHGFCASVFWHRNIPELAKKYKVVALDLGFGWSDKALIEYDAMVVDQIDFMK EVLHGFAPVVGSLGGTALSVAGLPEQVTGVALNSAQPAESKRKEADETVIT KPIVKPLKEIFQRVVLGFLFWAQKQPSRIEVLKSVYIDSTNVDDYLVESISKPAFD NAGEVYRLMTRFLTNQSRYTLDLSVLSKMTCPLLLVGDLDPWGPAPAKAEIKAFYSN SSLVHLQAGCHPDEVEPAVNVKALLDWSLINSIASKPASPVLLET"

CDS
1..1182
/gene="unannotated coding sequence from BAC C1C7A10"
217..1182
/gene="unannotated coding sequence from BAC C1C7A10"
/codon_start=1
/product="Unknown protein"
/protein_id="AAK43858.1"
/db_xref="GI:13877561"
/translation="MDTSSASQSVQSERLAFKPEGYNFWWRGKHIVYVQVGGESPL LVLIHGFASVFWHRNIPELAKKYKVVALDLGFGWSDKALIEYDAMVVDQIDFM KEVYKGFAPVVGSLGGTALSVAGLPEQVTGVALNSAQPAESKRKEADETVIT KPIVKPLKEIFQRVVLGFLFWAQKQPSRIEVLKSVYIDSTNVDDYLVESISKPAFD NAGEVYRLMTRFLTNQSRYTLDLSVLSKMTCPLLLVGDLDPWGPAPAKAEIKAFYSN NSSLVHLQAGCHPDEVEPAVNVKALLDWSLINSIASKPASPVLLET"

ORIGIN
Query Match 91.1%; Score 16.4; DB 8; Length 1301;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 573 ACGGCATCGTCAGTTGCG 590
|||||

RESULT 20
AF370481
LOCUS Arabidopsis thaliana Unknown protein mRNA, complete cds.
DEFINITION Arabidopsis thaliana 1307 bp mRNA linear PLN 30-APR-2001

AF370481 GI:13877560
VERSION FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1307)
Lam, B., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.

TITLE Direct Submission
JOURNAL Submitted (17-APR-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT e-mail for correspondence: arabesequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Lam, B., Karlin-Neumann, G., Nguyen, M., Southwick, A., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Sakurai, T., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Lam, B., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source
Location/Qualifiers
1..1307
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="IV"
/clone="RAFL11-06-K14"
/ecotype="Columbia"
/note="This clone is in pBluescript"
1..1307
/gene="unannotated coding sequence from BAC C1C7A10"
217..1182
/gene="unannotated coding sequence from BAC C1C7A10"
/codon_start=1
/product="Unknown protein"
/protein_id="AAK43858.1"
/db_xref="GI:13877561"
/translation="MDTSSASQSVQSERLAFKPEGYNFWWRGKHIVYVQVGGESPL LVLIHGFASVFWHRNIPELAKKYKVVALDLGFGWSDKALIEYDAMVVDQIDFM KEVYKGFAPVVGSLGGTALSVAGLPEQVTGVALNSAQPAESKRKEADETVIT KPIVKPLKEIFQRVVLGFLFWAQKQPSRIEVLKSVYIDSTNVDDYLVESISKPAFD NAGEVYRLMTRFLTNQSRYTLDLSVLSKMTCPLLLVGDLDPWGPAPAKAEIKAFYSN NSSLVHLQAGCHPDEVEPAVNVKALLDWSLINSIASKPASPVLLET"

ORIGIN
Query Match 91.1%; Score 16.4; DB 8; Length 1307;
Best Local Similarity 94.4%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 577 ACGGCATCGTCAGTTGCG 594
|||||

RESULT 21
SCYBR145W/c

```

LOCUS          SCYBR145W          1710 bp      DNA          linear          PLN 11-AUG-1997
DEFINITION     S.cerevisiae chromosome II reading frame ORF YBR145W.
ACCESSION      Z36014.1 Y13134
VERSION        Z36014.1 GI:536447
KEYWORDS
SOURCE
ORGANISM       Saccharomyces cerevisiae (baker's yeast)
REFERENCE
AUTHORS        Feldmann,H., Aigle,M., Aljinovic,G., Andre,B., Baclet,M.C.,
                Barthe,C., Baur,A., Becam,A.M., Biteau,N., Boles,E., Brandt,T.,
                Brendel,M., Bruckner,M., Bussereau,F., Christiansen,C.,
                Contreras,R., Crouzet,M., Cziepluch,C., Demolis,N., Delaveau,T.,
                Doignon,F., Dondoy,H., Dueterhus,S., Dubois,E., Dujon,B., El
                Bakoury,M., Entian,K.D., Feuermann,M., Fiers,W., Fobo,G.M.,
                Fritz,C., Gassenhuber,H., Glansdorff,N., Goffeau,A., Grivell,L.A.,
                de Haan,M., Hein,C., Herbert,C.J., Hollenberg,C.P., Holmstrom,K.,
                Jacq,C., Jacquet,M., Jauniaux,J.C., Jonniaux,J.L., Kallesoe,T.,
                Kiesau,P., Kirchthath,L., Koetter,P., Korol,S., Liebl,S., Logghe,M.,
                Lohan,A.J.E., Louis,E.J., Li,Z.Y., Maat,M.J., Mallet,L.,
                Mannhaupt,G., Messenguy,F., Miosga,T., Molemans,F., Mueller,S.,
                Nasir,F., Obermaier,B., Perea,J., Pierard,A., Piravandi,E.,
                Pohl,F.M., Pohl,T.M., Potier,S., Proft,M., Purnelle,B., Ramezani
                Rad,M., Rieger,M., Rose,M., Schaaff-Gerstenschlaeger,I.,
                Scherens,B., Schwarzlöse,C., Skala,J., Slonimski,P.P.,
                Smits,P.H.M., Souciet,J.L., Steensma,H.Y., Stucka,R.,
                Urrestazu,A., van der Aart,Q.J., van Dyck,L., Vassarotti,A.,
                Vetter,I., Vierendeels,F., Vissers,S., Wagner,G., de Wergifosse,P.,
                Wolfe,K.H., Zagulski,M., Zimmermann,F.K., Mewes,H.W. and Kleine,K.
                Complete DNA sequence of yeast chromosome II
                ENBO J. 13 (24), 5795-5809 (1994)
                95112788
                7813418
                2 (bases 1 to 1710)
                Entian,K.D., Koetter,P., Rose,M., Becker,J., Grey,M., Li,Z.,
                Niegemann,E., Schenk-Groeninger,R., Servos,J., Wehner,E.,
                Wolter,R., Brendel,M., Bauer,J., Braun,H., Dern,K., Dueterhus,S.,
                Gruenbein,R., Hedges,D., Kiesau,P., Korol,S., Krebs,B., Proft,M.,
                Slegers,K., Baur,A., Boles,E., Miosga,T.,
                Schaaff-Gerstenschlaeger,I. and Zimmermann,F.K.
                Unpublished
                3 (bases 1 to 1710)
                MIPS.
                Direct Submission
                Submitted (30-AUG-1994) Data collected by MIPS on behalf of the
                European yeast chromosome II sequencing project. MIPS at the
                Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
                Martinsried, FRG; E-mail: Mewes@mips.embnat.org
FEATURES
source          Location/Qualifiers
                1. .1710
                   /organism="Saccharomyces cerevisiae"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:4932"
                   /chromosome="II"
                213. .1268
                   /gene="ADH5"
                213. .1268
                   /gene="ADH5"
                   /note="ORF YBR145W"
                   /codon_start=1
                   /protein_id="CAA85103.1"
                   /db_xref="GI:536448"
                   /db_xref="GOA:P38113"
                   /db_xref="SGD:S0000349"
                   /translacion="MPSQVPEKQKAVFYETDGLKLYKDVTPPEKPNILVHVKYS
                   GVCHSDLAHGGWPFLKPLGLEGHAGVWVKLSNVGKWKGFAGIKWLNGTCM
                   SCEYCEVGNESQCPLYDGTGTHDTCFQEVATADAVQAHPNNVLAEPILCAGI
                   TVYKALKRANVIPQGWITISGACGGLGSLAIQYALANGRYVIGIDGNKRRKFLQEG
                   GEIFDTEKDIIVGAIKATNGSHGVINVSAAIEASTYRCRPNGTVLVGMFA
                   HAYCNSDVFNQVVKVSIIGVSCVGNRDTREALDFFARGLIKPSIHLAGLSDVPEIFA
                   KMEKEIVGVYVETS"
gene
CDS

```

```

ORIGIN
Query Match          91.1%;      Score 16.4;   DB 8;   Length 1710;
Best Local Similarity 94.4%;   Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTCCG 18
        |||||
        625  ACGGCATCGCAGTTCCG 608

Db

RESULT 22
ATAP22
LOCUS     ATAP22          198354 bp      DNA          linear          PLN 07-APR-1999
DEFINITION Arabidopsis thaliana DNA chromosome 4, ESSA I AP2 contig fragment
                No. 2.
ACCESSION  Z99708
VERSION    Z99708.1 GI:4006885
KEYWORDS
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE  1
            Bevan,M., Terry,N., Vos,P., Heijnen,L., Mewes,H.W., Mayer,K.F.X.
            and Schueller,C.
            Unpublished
            2 (bases 1 to 198354)
            EU Arabidopsis sequencing,project.
            Direct Submission
            Submitted (07-APR-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuellemips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
            On Dec 12, 1998 this sequence version replaced gi:2464894.
            The annotation of this entry was produced with considerable
            contributions from Stephane Rombauts and Pierre Rouze, Department
            of Genetics, University of Ghent, Ledengankstraat 35, 9000 Ghent,
            BE, E-mail: strom@genrug.ac.be, piro@genrug.ac.be
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATAP21 at the 5' end.
FEATURES
source          Location/Qualifiers
                1. .198354
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /variety="Columbia"
                   /db_xref="taxon:3702"
                   /chromosome="4"
                1. .10046
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:3702"
                   /clone="cosmid CIC7A10.4E12"
                18152. .35497
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:3702"
                   /clone="cosmid CIC7A10.5C09"
                35444. .55957
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:3702"
                   /clone="cosmid CIC7A10.2H02"
                55964. .76032
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:3702"
                   /clone="cosmid CIC7A10.4B03"

```


exon 8796..8921
/gene="AT4g36190"
/number=13
intron 8922..9011
/gene="AT4g36190"
/number=13
exon 9012..9185
/gene="AT4g36190"
/number=14
intron 9186..9264
/gene="AT4g36190"
/number=14
exon 9265..9450
/gene="AT4g36190"
/number=15
intron 9451..9551
/gene="AT4g36190"
/number=15
exon 9552..9668
/gene="AT4g36190"
/number=16
intron 9669..9767
/gene="AT4g36190"
/number=16
exon 9768..9821
/gene="AT4g36190"
/number=17
intron 9822..9900
/gene="AT4g36190"
/number=17
exon 9901..9972
/gene="AT4g36190"
/number=18
intron 9973..10582
/gene="AT4g36190"
/number=18
exon 10583..10657
/gene="AT4g36190"
/number=19
intron 10658..10726
/gene="AT4g36190"
/number=19
exon 10727..10865
/gene="AT4g36190"
/number=20
intron 10866..10979
/gene="AT4g36190"
/number=20
exon 10980..11095
/gene="AT4g36190"
/number=21
tRNA 11360..11432
/product="tRNA-Glu"
/note="codon recognized: GAG"
gene 11536..11982
/gene="AT4g36200"
complement(11536..11982)
gene /gene="AT4g36200"
CDS complement(11536..11982)
/gene="AT4g36200"

Query Match 91.1%; Score 16.4; DB 8; Length 198750;
Best Local Similarity 94.4%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 122199 ACGGCATGTCAGTTGCG 122182

RESULT 24
AE017234 313846 bp DNA linear BCT 02-FEB-2004
LOCUS AE017234

DEFINITION Mycobacterium avium subsp. paratuberculosis str. k10, section 8 of
16 of the complete genome.
ACCESSION AE017234 AE016958
VERSION AE017234.1 GI:41396359
KEYWORDS
SOURCE
ORGANISM
Mycobacterium avium subsp. paratuberculosis str. k10
Mycobacterium avium subsp. paratuberculosis str. k10
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
avium complex (MAC).
REFERENCE 1 (bases 1 to 313846)
AUTHORS Li, L., Bannantine, J., Zhang, Q., Amonsin, A., Alt, D. and Kapur, V.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2003) Biomedical Genomics Center, University of
Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA
FEATURES
source
1..313846
/organism="Mycobacterium avium subsp. paratuberculosis
str. k10"
/mol_type="genomic DNA"
/strain="k10"
/sub_species="paratuberculosis"
/db_xref="taxon:262316"
complement(238..639)
/locus_tag="MAP1907c"
complement(238..639)
/locus_tag="MAP1907c"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAS04224.1"
/db_xref="GI:41396360"
/translation="MPLSDHEQRLDQIESALYAEADPKFASSVREGGRFAPTAARRLLQ
GVALFVIGLAMLVSGVAFKATMIGNFILSVFGIVMFGGVFAITGPRLPGRPDLSG
PAPTSLRQRENRRSGGSFTSRMEDRFRFRDD"
1019..1621
/locus_tag="MAP1908"
1019..1621
/locus_tag="MAP1908"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAS04225.1"
/db_xref="GI:41396361"
/translation="MAIFLDLPHDMERLGDALAYVVDAMRYPRDTENORAAWMLE
HIRRGWQGAIVAEVDDDDQMPFAEELASAPLLGYGYPGAPGQWQQQVVLGLQ
RGSPPOEIAIWMNDYFELTELHHPRAOQGRGLCEALARLLAGRAEKVLLSTPETN
GEFNARWRLYRLGLFTDIIRRYHFAGDPRFPAILGRELP"
1690..2421
/gene="lppm"
/locus_tag="MAP1909"
1690..2421
/gene="lppm"
/locus_tag="MAP1909"
/codon_start=1
/transl_table=1
/product="lppm"
/protein_id="AAS04226.1"
/db_xref="GI:41396362"
/translation="MRARPAPRRRSGQAARMRRRLAFAMLLLVVPLATGCLRVA
SLTISDDLVSGEIVAAKPKTKDTGTPQLDNNLPFSQKVAVSNYSDGVGQAVP
SDLTFAELPOLANNSDAGVLSLRNGNLVLEGRADLTSLTDPADVELTVAFPG
VVTSTGDRVEPDVWSWKLKPGVSVTMTAQARYTDPNTRSFTHAVVWLGIAFAAAGV
VGLLAWYSDRSRPLTAPDQPPPE"
complement(2425..3333)
/locus_tag="MAP1910c"
complement(2425..3333)
/locus_tag="MAP1910c"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAS04227.1"
/db_xref="GI:41396363"


```

REFERENCE
AUTHORS
  Bradyrhizobiaceae: Rhodopseudomonas.
  1 (bases 1 to 349746)
  Larimer, F.W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L.,
  Land, M.L., Pelletier, D.A., Beatty, J.T., Lang, A.S., Tabita, F.R.,
  Gibson, J.L., Hanson, T.E., Bobat, C., Torres, J.L., Peres, C.,
  Harrison, P.H., Gibson, J., and Harwood, C.S.
  Complete genome sequence of the metabolically versatile
  photosynthetic bacterium Rhodopseudomonas palustris
  Nat. Biotechnol. 22 (11), 55-61 (2004)
  14704707

JOURNAL
PUBMED
  2 (bases 1 to 349746)
  Larimer, F.W. and Harwood, C.S.
  Rhodopseudomonas genome consortium
  Direct Submission
  Submitted (24-JUL-2003) Submitted on behalf of the Rhodopseudomonas
  genome consortium, the DOE Joint Genome Institute, Production
  Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
  USA, and the Genome Analysis Group, Oak Ridge National Laboratory,
  1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
  larimerf@ornl.gov

FEATURES
  source
    1. .349746
    /organism="Rhodopseudomonas palustris CGA009"
    /mol_type="genomic DNA"
    /strain="CGA009"
    /db_xref="taxon:258594"
    96..521
    /locus_tag="RPA3709"
    96..521
    /locus_tag="RPA3709"
    /locus_tag="RPA3709"
    /function="InterPro IPR000971 COGs COG1017"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="possible hemoprotein"
    /protein_id="CAE29150.1"
    /db_xref="GI:39650628"
    /translation="MTPQAIYVQSFARLTPISDETATLTYERLFATAPELRPLFH
    GDMRQRKGLMSTUNVVVGLTDTITLPTATGRLARLHVAYGVAHYTPFGAALLWT
    LERELGAEWTPELAAWRDAYAMLSLTAAYGDTTPS"
    564..2336
    /gene="nirA"
    /locus_tag="RPA3710"
    564..2336
    /gene="nirA"
    /locus_tag="RPA3710"
    /locus_tag="RPA3710"
    /EC_number="1.7.1.1"
    /function="InterPro
    IPR001424:IPR005117:IPR006066:IPR006067 COGs COG0155"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="possible ferredoxin-nitrite reductase"
    /protein_id="CAE29151.1"
    /db_xref="GI:39650629"
    /translation="MGSEFSIDQRYLEGATGTAAARAGSLPAGSAASPSGPDAI
    HHQADRVATAGKLNDOERKQHFDFAYARLKQAKNAAPKADNFRWRYGLF
    YVAPQTSYMCRLIPNGILTSWQAGLVADLTADGYAHVTTFRANLQMEIAPKDA
    VLLQGLIESGLMARSGADNIRNVTSATAGIDPDLIDTRAVARDWHFYLINERAI
    FGLPRKFNVAFDGGGLIPTLEDNTDIAFQAVTIGDGHGVAQGVYFRALGGLTGHKDF
    ARDTGVIVADEATEVADAIVRVFIEHGDRTDRNKARKLYVDLAFGFDKYLELVEDKL
    GRKLIRVPAEAAQDRPAQDRAAHGLGVHPQQAQGNWIGVRLATGHLTSVQMRGLAEIA
    RSFGDGLRLTYWNQLISGVDPDARVDTAAIEALGLAITASPLRAGLIACGASGC
    RFAAAHKTETAEIAQYCEPRVPLDTPINHLTGCHNSCAQHFISDGLICAKVAVSD
    EDSVEGFHIVHGQGVGDAIAGTVELRVKQDDAPRVIERMLQTYLARRASPAETFLA
    FVRRODAALQQLFAYEQSPETVA"
    2333..3952
    /locus_tag="RPA3711"
    2333..3952
    /locus_tag="RPA3711"
    /EC_number="1.8.1.2"
    /function="InterPro IPR001433:IPR001709:IPR003097 COGs
    COG0363"

gene
  CDS
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="possible sulfite reductase (NADPH)"
    /protein_id="CAE29152.1"
    /db_xref="GI:39650630"
    /translation="MSONLOPPIDLLVPETAPFSEORAWLNGPAGLVSIDGCGVTA
    LSPQAAALLAGPAPVAEDDDGGAPHDQSLPIGERHQLAEGKPLRWKLMAAQQDD
    CGCGYDCRNYSGAIFDQETRLNLCAPGGKDTARMVKTLTEQIGAAPTLNAPSLAA
    DAEFVALPPRGTSRDNPATKLSRKLNKPGESEKTMHVEFEELDCLSYAGDSFG
    LFTPDNALVDLHALGAPAEFFIAQSLTLQRLDLSVSPADPMLPQLIYSITGGD
    KRKARSLASGEDPDGAATLDVLAALKFPLGRPDPEAFTEALDPLQRLYSISSP
    KTTPTGRLSLTVDCVRYAIGKQRIGVCSTGLAERVRGSETLRYVQKAAHHALPSDPN
    OPTIMIGPTGIAEFRAFLHERQAIQAPKRWLFFGHQRSASDFEYDELKAMKTAGH
    LTRUTLAWSRDSGSKIYVQDRIRREVRLWSLAEAGSLYVCGDAKRWAKDVERALVD
    IVAQHGSARSAEATAFVSELKQKQGYQDDVY"
    complement (4092..4863)
    /gene="pimE pseudogene"
    /locus_tag="RPA3712"
    /pseudo
    complement (join (4092..4487,4486..4863))
    /gene="pimE pseudogene"
    /locus_tag="RPA3712"
    /EC_number="1.1.1.-"
    /function="InterPro IPR002198 COGs COG1028"
    /note="frameshift"
    /pseudo
    /codon_start=1
    /evidence=experimental
    /transl_table=11
    /product="short chain dehydrogenase"
    /db_xref="PSEUDO:CAE29153.1"
    complement (5013..6155)
    /gene="pimD"
    /locus_tag="RPA3713"
    complement (5013..6155)
    /gene="pimD"
    /locus_tag="RPA3713"
    /EC_number="1.3.1.62"
    /function="InterPro IPR001687:IPR006090:IPR006091 COGs
    COG1960"
    /note="observed by proteomics
    Citation: Proteomics from Verberkmoes et al. (2003)
    unpublished"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="pimeloyl-CoA dehydrogenase (small subunit)"
    /protein_id="CAE29154.1"
    /db_xref="GI:39650631"
    /translation="MDFLSEERQLKSEVGLKGSYDFDSRKKYAKEKGGWRAVW
    GKFAEQGLGLPFSEEDGFGAGAVETMIVMEALGHSVLVLEFYLPTVVGGLFRRAG
    SAAQAAHLPGLIIGDSKTFAPQAEKNSRWLDGVDVTSKAGSGWIDGSKFVVLNG
    EAADLTIVTARTKGQRDRDTGVGVFLVPADAKGITRKGYPQDGLHAADITFTGVQVG
    ADAALIGFENALIELIAVDDARTALCAEAVGLMDESLTTVEIKTRKQGVPIGSF
    QVLOHRAADNFVATEQARSMAFATMAAEFFDDAKERAGATAAAKQVIGKSKFVCGQS
    IQLHGGLTGMTMEAKIGHYFKRLTMEIQTFGDTDHLLARVSAGGLI"
    complement (6324..7514)
    /gene="pimC"
    /locus_tag="RPA3714"
    complement (6324..7514)
    /gene="pimC"
    /locus_tag="RPA3714"
    /EC_number="1.3.1.62"
    /function="InterPro IPR006090:IPR006091:IPR006092 COGs
    COG1960"
    /note="observed by proteomics
    Citation: Proteomics from Verberkmoes et al. (2003)
    unpublished"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="pimeloyl-CoA dehydrogenase (large subunit)"

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 69.9409 Seconds
(without alignment)
1190.710 Million cell updates/sec

Title: US-09-674-277-18

Perfect score: 14

Sequence: 1 ggcacgcctggtg 14

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14	100.0	774	9	US-09-738-626-920
2	14	100.0	1285	17	US-10-282-122A-36930, A
3	14	100.0	1551	17	US-10-282-122A-40133
4	14	100.0	3309400	9	US-09-738-626-1
5	13	92.9	25	19	US-10-719-900-375244
6	13	92.9	60	10	US-09-908-975-15299
C 7	13	92.9	132	9	US-09-864-761-18565
C 8	13	92.9	380	17	US-10-282-122A-22906
C 9	13	92.9	397	10	US-09-918-995-17778
C 10	13	92.9	414	18	US-10-437-963-83252
C 11	13	92.9	445	17	US-10-424-599-75081

C 12	13	92.9	448	9	US-09-864-761-1814
C 13	13	92.9	464	18	US-10-853-047-3281
C 14	13	92.9	589	18	US-10-021-323-3974
15	13	92.9	637	13	US-10-027-632-201074
16	13	92.9	637	13	US-10-027-632-201075
17	13	92.9	637	17	US-10-027-632-201074
18	13	92.9	637	17	US-10-027-632-201075
C 19	13	92.9	695	18	US-10-425-115-143802
C 20	13	92.9	781	18	US-10-425-115-65151
C 21	13	92.9	906	18	US-10-723-860-1431
C 22	13	92.9	915	13	US-10-027-632-323949
C 23	13	92.9	915	17	US-10-027-632-323949
C 24	13	92.9	960	9	US-09-815-242-6170
C 25	13	92.9	960	17	US-10-282-122A-20469
C 26	13	92.9	1101	18	US-10-425-115-12868
C 27	13	92.9	1315	18	US-10-283-975A-266
C 28	13	92.9	1407	17	US-10-369-493-44633
C 29	13	92.9	1461	17	US-10-282-122A-14037
C 30	13	92.9	1467	19	US-10-838-834-15
C 31	13	92.9	1481	18	US-10-620-669-12
C 32	13	92.9	1608	18	US-10-437-963-43652
C 33	13	92.9	1728	17	US-10-287-971-45
C 34	13	92.9	1876	18	US-10-425-115-65153
C 35	13	92.9	1986	17	US-10-168-417A-1
C 36	13	92.9	2115	17	US-10-136-145-1
C 37	13	92.9	2130	9	US-09-862-260A-1
C 38	13	92.9	2130	9	US-09-812-238B-1
C 39	13	92.9	2130	15	US-10-207-655-76
C 40	13	92.9	2131	13	US-10-047-539-3
C 41	13	92.9	2172	10	US-09-898-860-26
C 42	13	92.9	2172	18	US-10-685-977-26
C 43	13	92.9	2463	17	US-10-282-122A-8333
C 44	13	92.9	2586	17	US-10-369-493-43449
C 45	13	92.9	2685	18	US-10-425-115-65144

ALIGNMENTS

RESULT 1

US-09-738-626-920/c
; Sequence 920, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 920
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-920

Query Match

100.0%; Score 14; DB 9; Length 774;

Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 420 GGCATCGTCAGTTG 407

RESULT 2

US-10-282-122A-36930
; Sequence 36930, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36930
; LENGTH: 1285
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A

US-10-282-122A-36930

Query Match 100.0%; Score 14; DB 17; Length 1285;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 403 GGCATCGTCAGTTG 416

RESULT 3

US-10-282-122A-40133
; Sequence 40133, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40133
; LENGTH: 1551
; TYPE: DNA
; ORGANISM: Salmonella typhi

US-10-282-122A-40133

Query Match 100.0%; Score 14; DB 17; Length 1551;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 694 GGCATCGTCAGTTG 707

RESULT 4

US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOHI, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 100.0%; Score 14; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
Db 870272 GGCATCGTCAGTTG 870285
|||||

RESULT 5

US-10-719-900-375244
; Sequence 375244, Application US/10719900
; Publication No. US2005002616441
; GENERAL INFORMATION:

; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 375244
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-375244

Query Match 92.9%; Score 13; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
Db 5 GGCATCGTCAGTT 17
|||||

RESULT 6

US-09-908-975-15299/c
; Sequence 15299, Application US/09908975
; Publication No. US200301658431
; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15299

; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-15299

Query Match 92.9%; Score 13; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
Db 16 GCATCGTCAGTTG 4
|||||

RESULT 7

US-09-864-761-18565/c
; Sequence 18565, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18565
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007078.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.3

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
; OTHER INFORMATION: NT HIT: g16005983, EVALUE 1.00e-61
; OTHER INFORMATION: SWISSPROT HIT: P53785, EVALUE 6.00e-18
; OTHER INFORMATION: EST_HUMAN HIT: BE747227.1, EVALUE 2.00e-61
US-09-864-761-18565

Query Match 92.9%; Score 13; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 15 GGCATCGTCAGTT 3

RESULT 8
US-10-282-122A-22906/c
; Sequence 22906, Application US/10282122A
; Publication No. US20040029129A1

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; PRIOR FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 22906

; LENGTH: 380

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-10-282-122A-22906

Query Match 92.9%; Score 13; DB 17; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 63 GGCATCGTCAGTT 51

RESULT 9

US-09-918-995-17778/c

; Sequence 17778, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 17778

; LENGTH: 397

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)..(397)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-17778

Query Match 92.9%; Score 13; DB 10; Length 397;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 392 GGCATCGTCAGTT 380

RESULT 10

US-10-437-963-83252/c

; Sequence 83252, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 83252

; LENGTH: 414

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_82601C.1

US-10-437-963-83252

Query Match 92.9%; Score 13; DB 18; Length 414;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
|||||

Db 391 GCATCGTCAGTTG 379

RESULT 11

US-10-424-599-75081/c
; Sequence 75081, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 75081
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_38811C.1
US-10-424-599-75081

Query Match 92.9%; Score 13; DB 17; Length 445;

Best Local Similarity 100.0%; Pred. No. 1.6e+03; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14

Db 150 GCATCGTCAGTTG 138

RESULT 12

US-09-864-761-1814/c
; Sequence 1814, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1814
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007078.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.7
US-09-864-761-1814

Query Match 92.9%; Score 13; DB 9; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.6e+03; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13

Db 294 GGCATCGTCAGTT 282

RESULT 13

US-10-653-047-3281/c
; Sequence 3281, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3281
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(464)

```
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3281

Query Match          92.9%; Score 13; DB 18; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCATCGTCAGTTG 14
Db 422 GCATCGTCAGTTG 410
|||||

RESULT 14
US-10-021-323-3974/c
; Sequence 3974, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 3974
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(589)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3825-030-Q6-K6-E1
US-10-021-323-3974

Query Match          92.9%; Score 13; DB 18; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 52 GGCATCGTCAGTT 40
|||||

RESULT 15
US-10-027-632-201074
; Sequence 201074, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201075
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201075

Query Match          92.9%; Score 13; DB 13; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 388 GGCATCGTCAGTT 400
|||||

RESULT 16
US-10-027-632-201075
; Sequence 201075, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201075
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201075

Query Match          92.9%; Score 13; DB 13; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 388 GGCATCGTCAGTT 400
|||||

RESULT 17
US-10-027-632-201074
; Sequence 201074, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```

```
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-3281

Query Match          92.9%; Score 13; DB 18; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCATCGTCAGTTG 14
Db 422 GCATCGTCAGTTG 410
|||||

RESULT 14
US-10-021-323-3974/c
; Sequence 3974, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 3974
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(589)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3825-030-Q6-K6-E1
US-10-021-323-3974

Query Match          92.9%; Score 13; DB 18; Length 589;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 52 GGCATCGTCAGTT 40
|||||

RESULT 15
US-10-027-632-201074
; Sequence 201074, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201075
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201075

Query Match          92.9%; Score 13; DB 13; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 388 GGCATCGTCAGTT 400
|||||

RESULT 16
US-10-027-632-201075
; Sequence 201075, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201075
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201075

Query Match          92.9%; Score 13; DB 13; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 388 GGCATCGTCAGTT 400
|||||

RESULT 17
US-10-027-632-201074
; Sequence 201074, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201074
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201074

Query Match 92.9%; Score 13; DB 17; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 388 GGCATCGTCAGTT 400

RESULT 18
US-10-027-632-201075
; Sequence 201075, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201075
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201075

Query Match 92.9%; Score 13; DB 17; Length 637;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 388 GGCATCGTCAGTT 400

RESULT 19
US-10-425-115-143802/c

; Sequence 143802, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 143802
; LENGTH: 695
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(695)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_62625C.1
US-10-425-115-143802

Query Match 92.9%; Score 13; DB 18; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
|||||
Db 683 GCATCGTCAGTTG 671

RESULT 20
US-10-425-115-65151
; Sequence 65151, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 65151
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(781)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_159414C.1
US-10-425-115-65151

Query Match 92.9%; Score 13; DB 18; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
|||||
Db 314 GCATCGTCAGTTG 326

RESULT 21
US-10-723-860-1431/c
; Sequence 1431, Application US/10723860

```
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1431
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1431

Query Match          92.9%; Score 13; DB 18; Length 906;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGCATCGTCAGTT 13
Db      422 GGCATCGTCAGTT 410

RESULT 22
US-10-027-632-323949/c
; Sequence 323949, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323949
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323949

Query Match          92.9%; Score 13; DB 13; Length 915;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  GCATCGTCAGTTG 14
Db      27  GCATCGTCAGTTG 15

RESULT 23
US-10-027-632-323949/c
; Sequence 323949, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323949
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323949

Query Match          92.9%; Score 13; DB 13; Length 915;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  GCATCGTCAGTTG 14
Db      27  GCATCGTCAGTTG 15

RESULT 24
US-09-815-242-6170/c
; Sequence 6170, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; Sequence 323949, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 323949
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-323949

Query Match          92.9%; Score 13; DB 17; Length 915;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  GCATCGTCAGTTG 14
Db      27  GCATCGTCAGTTG 15

RESULT 24
US-09-815-242-6170/c
; Sequence 6170, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6170
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(960)
US-09-815-242-6170

Query Match 92.9%; Score 13; DB 9; Length 960;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
| | | | | | | | | |
Db 172 GCATCGTCAGTTG 160

RESULT 25

US-10-282-122A-20469/c
; Sequence 20469, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20469
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-282-122A-20469

Query Match 92.9%; Score 13; DB 17; Length 960;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
| | | | | | | | | |
Db 172 GCATCGTCAGTTG 160

Search completed: March 12, 2005, 00:25:24
Job time : 73.9409 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:06:25 ; Search time 78.4641 Seconds
(without alignments)
8645.645 Million cell updates/sec

Title: US-09-674-277-18
Perfect score: 14
Sequence: 1 ggcacgtcagttg 14

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	100.0	14	6	BD205228 Nucleotid
2	14	100.0	14	6	AX011314 Sequence
3	14	100.0	16	6	BD205229 Nucleotid
4	14	100.0	16	6	AX011315 Sequence
5	14	100.0	18	6	BD205230 Nucleotid
6	14	100.0	18	6	AX011316 Sequence
7	14	100.0	31	6	BD205222 Nucleotid
8	14	100.0	31	6	BD205223 Nucleotid
9	14	100.0	31	6	AX011308 Sequence
10	14	100.0	31	6	AX011309 Sequence
11	14	100.0	431	3	AF353336 Vargula h
12	14	100.0	443	3	AF353331 Vargula h
13	14	100.0	536	1	UBA319838 Unculture
14	14	100.0	774	6	BD163121 Novel pol
15	14	100.0	774	6	AX121004 Sequence
16	14	100.0	1489	6	BD205211 Nucleotid
17	14	100.0	1489	6	AX011297 Sequence
18	14	100.0	1517	1	PSIS801 X57269 P. syringae
19	14	100.0	2055	4	AB022426 Sus scrofa

C 20	14	100.0	5145	6	AX073964 Sequence
C 21	14	100.0	5916	3	AF195498 Drosophil
C 22	14	100.0	5938	3	AF250842 Drosophil
C 23	14	100.0	5959	3	AB031048 Drosophil
C 24	14	100.0	5962	3	AY069579 Drosophil
C 25	14	100.0	22694	1	AE008914 Salmonell
C 26	14	100.0	39261	2	AC020271 Drosophil
C 27	14	100.0	40110	1	AY603980 Pseudomon
C 28	14	100.0	40984	2	AC136807 Rattus no
C 29	14	100.0	79733	2	AC006579 Drosophil
C 30	14	100.0	87197	2	AC014071 Drosophil
C 31	14	100.0	88037	1	AL627284 Salmonell
C 32	14	100.0	92077	1	AF074613 Escherich
C 33	14	100.0	92077	6	AX191727 Sequence
C 34	14	100.0	92721	1	AB011549 Escherich
C 35	14	100.0	92721	6	AX191725 Sequence
C 36	14	100.0	110000	1	BX571965_01 Continuation (2 of
C 37	14	100.0	110000	1	CP000010_01 Continuation (2 of
C 38	14	100.0	154681	3	AC095008 Drosophil
C 39	14	100.0	162689	2	AC117145 Rattus no
C 40	14	100.0	163098	3	AC010055 Drosophil
C 41	14	100.0	180555	9	AC079384 Homo sapi
C 42	14	100.0	182960	3	AC010017 Drosophil
C 43	14	100.0	183161	3	AC007851 Drosophil
C 44	14	100.0	183925	2	AC078870 Homo sapi
C 45	14	100.0	195205	2	AC131416 Rattus no

ALIGNMENTS

RESULT 1	BD205228	14 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD205228				
DEFINITION	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).				
ACCESSION	BD205228				
VERSION	BD205228.1	GI:33014998			
KEYWORDS	JP 2002512813-A/18.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 14)				
AUTHORS	Frechon,D.T.M., Laure,F.C. and Thierry,D.				
TITLE	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)				
JOURNAL	Patent: JP 2002512813-A 18 08-MAY-2002;				
COMMENT	BIORAD PASTEUR				
	OS 'Unidentified				
	PN JP 2002512813-A/18				
	PD 08-MAY-2002				
	PF 27-APR-1998 FR 98/05329				
	PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDE LAURE, PI				
	DOMINIQUE THIERRY				
	PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC				
	Strandedness: Single;				
	CC Topology: Linear;				
	CC Nucleotide sequence for detecting enterohemorrhagic CC				
	Escherichia coli				
	CC (EHEC).				
FEATURES	FT Key	Location/Qualifiers			
	1..14	1..14			
	FT source	/organism='Unidentified'.			
	source	1..14			
		/organism='unidentified'			
		/mol_type='genomic DNA'			
		/db_xref='taxon:32644'			
ORIGIN					
	Query Match	100.0%;	Score 14;	DB 6;	Length 14;
	Best Local Similarity	100.0%;	Pred. No. 1.7e+03;		

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 1 GGCATCGTCAGTTG 14

RESULT 2
AX011314
LOCUS
DEFINITION Sequence 18 from Patent WO9955908. 14 bp DNA linear PAT 06-SEP-2000
ACCESSION AX011314
VERSION AX011314.1 GI:9997864
KEYWORDS
SOURCE
ORGANISM Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Thierry,D., Frechon,D.T. and Laure,F.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehec)

JOURNAL Patent: WO 9955908-A 18 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDEINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES
source Location/Qualifiers
1.14
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN

Query Match 100.0%; Score 14; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 1 GGCATCGTCAGTTG 14

RESULT 3
BD205229
LOCUS
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC). 16 bp DNA linear PAT 17-JUL-2003
ACCESSION BD205229
VERSION BD205229.1 GI:33014999
KEYWORDS JP 2002512813-A/19.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 16)
AUTHORS Frechon,D.T.M., Laure,F.C. and Thierry,D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)

JOURNAL Patent: JP 2002512813-A 19 08-MAY-2002;
BIORAD PASTEUR

COMMENT OS Unidentified
PN JP 2002512813-A/19
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PI 28-APR-1998 FR 98/05329
PR DOMINIQUE THERESE MARIE FRECHON,FRANCOISE CLAUDEINE LAURE, PI DOMINIQUE THIERRY
PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC Escherichia coli
CC (EHEC).
FH Key Location/Qualifiers
FT source 1..16

FT Location/Qualifiers
source 1..16
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 14; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 2 GGCATCGTCAGTTG 15

RESULT 4
AX011315
LOCUS
DEFINITION Sequence 19 from Patent WO9955908. 16 bp DNA linear PAT 06-SEP-2000
ACCESSION AX011315
VERSION AX011315.1 GI:9997865
KEYWORDS
SOURCE
ORGANISM Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Thierry,D., Frechon,D.T. and Laure,F.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehec)

JOURNAL Patent: WO 9955908-A 19 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDEINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES
source Location/Qualifiers
1..16
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN

Query Match 100.0%; Score 14; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 2 GGCATCGTCAGTTG 15

RESULT 5
BD205230
LOCUS
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC). 18 bp DNA linear PAT 17-JUL-2003
ACCESSION BD205230
VERSION BD205230.1 GI:33015000
KEYWORDS JP 2002512813-A/20.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 18)
AUTHORS Frechon,D.T.M., Laure,F.C. and Thierry,D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)

JOURNAL Patent: JP 2002512813-A 20 08-MAY-2002;
BIORAD PASTEUR

COMMENT OS Unidentified
PN JP 2002512813-A/20
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329


```

PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08, C07K14/245, C12N1/21, C12N15/09, C12Q1/68, C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FH Key Location/Qualifiers
FT source 1..18
/organism="Unidentified".
FEATURES
source
1..18
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 14; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 3 GGCATCGTCAGTTG 16

RESULT 6
AX011316
LOCUS 18 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 20 from Patent WO955908.
ACCESSION AX011316
VERSION AX011316.1 GI:9997866
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1
AUTHORS Thierry, D., Frechon, D.T. and Laure, F.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehec)
JOURNAL Patent: WO 9955908-A 20 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANCOISE CLAUDINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES
source
1..18
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 14; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 3 GGCATCGTCAGTTG 16

RESULT 7
BD205222
LOCUS 31 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION BD205222
VERSION BD205222.1 GI:33014992
KEYWORDS JP 2002512813-A/12.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 31)

```

```

AUTHORS Frechon, D.T.M., Laure, F.C. and Thierry, D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 12 08-MAY-2002;
COMMENT BIORAD PASTEUR
OS Unidentified
PN JP 2002512813-A/12
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08, C07K14/245, C12N1/21, C12N15/09, C12Q1/68, C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FH Key Location/Qualifiers
FT source 1..31
/organism="Unidentified".
FEATURES
source
1..31
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 14; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 7 GGCATCGTCAGTTG 20

RESULT 8
BD205223
LOCUS 31 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION BD205223
VERSION BD205223.1 GI:33014993
KEYWORDS JP 2002512813-A/13.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 31)
AUTHORS Frechon, D.T.M., Laure, F.C. and Thierry, D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 13 08-MAY-2002;
COMMENT BIORAD PASTEUR
OS Unidentified
PN JP 2002512813-A/13
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08, C07K14/245, C12N1/21, C12N15/09, C12Q1/68, C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FH Key Location/Qualifiers
FT source 1..31
/organism="Unidentified".
FEATURES
source
1..31
/organism="unidentified"

```

32

AUTHORS Oakley, T.H. and Huber, D.R.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2001) Biology, Duke University, Box 90325,
Durham, NC 27708, USA

FEATURES
source
1. .443
/organism="Vargula hilgendorfi"
/mol_type="mRNA"
/db_xref="taxon:6674"
/clone="Vnp_S2_4_10"
/country="Japan"
/codon_start=2
/product="opsin"
/protein_id="AAU37506.1"
/db_xref="GI:17225632"
/translation="NVSRLNEDNKKMSAECKIAKVAMTNVFLMVAITWTYATVGMIG
VFQDSIITPLVSGIPLSVAKTASCFNPVVYAISHPRYQALQAMPWCVHEPTDDA
QSVASSTETTTSEKQ"
359. .443

CDS
3'UTR
ORIGIN
Query Match 100.0%; Score 14; DB 3; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 307 GGCATCGTCAGTTG 294

RESULT 13
UBA319838
LOCUS UBA319838 536 bp DNA linear BCT 27-SEP-2002
DEFINITION Uncultured bacterium SB-53-TW partial 16S rRNA gene.
ACCESSION AJ319838
VERSION AJ319838.1 GI:21489345
KEYWORDS 16S ribosomal RNA; 16S rRNA gene.
SOURCE uncultured bacterium
ORGANISM uncultured bacterium
Bacteria; environmental samples.

REFERENCE
1 Schaefer, H., Abbas, B., Witte, H. and Muyzer, G.
Genetic diversity of 'satellite' bacteria present in cultures of
marine diatoms
FEMS Microbiol. Ecol. 42 (1), 25-35 (2002)
Schaefer, H.
Direct Submission
Submitted (18-JUN-2001) Schaefer H., Biol. Oceanography,
Netherlands Institute for Sea Research, P.O. Box 59, NL-1790 AB Den
Burg (Texel), NETHERLANDS

FEATURES
source
1. .536
/organism="uncultured bacterium"
/mol_type="genomic DNA"
/db_xref="taxon:77133"
/clone="SB-53-TW"
/note="satellite bacterium of Thalassiosira weissflogii"
1. .536
/gene="16S rRNA"
<1. .536
/gene="16S rRNA"
/product="16S ribosomal RNA"

ORIGIN
Query Match 100.0%; Score 14; DB 1; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 226 GGCATCGTCAGTTG 239

RESULT 14
BD163121/c
LOCUS BD163121 774 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD163121
VERSION BD163121.1 GI:27868883
KEYWORDS JP 2002191370-A/920.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE
1 (bases 1 to 774)
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
Novel polynucleotide
Patent: JP 2002191370-A 920 09-JUL-2002;
KYOWA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/920
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
PC 04, C12P13/08,
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1:15),
PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
PC C12N15/00,
PC C12N5/00, C12N15/00
CC Novel polynucleotide
FH Key Location/Qualifiers
FT source 1. .774
/organism="Corynebacterium glutamicum".
Location/Qualifiers
1. .774
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 14; DB 6; Length 774;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 420 GGCATCGTCAGTTG 407

RESULT 15
AX121004/c
LOCUS AX121004 774 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 920 from Patent EP1108790.
ACCESSION AX121004
VERSION AX121004.1 GI:14037719
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
Novel polynucleotides
Patent: EP 1108790-A 920 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)

```

FEATURES             Location/Qualifiers
     source
     1..774
        /organism="Corynebacterium glutamicum"
        /mol_type="unassigned DNA"
        /db_xref="taxon:1718"

ORIGIN
Query Match          100.0%; Score 14; DB 6; Length 774;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 420 GGCATCGTCAGTTG 407

RESULT 16
BD205211             1489 bp      DNA      linear      PAT 17-JUL-2003
LOCUS
DEFINITION
Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION
BD205211
VERSION
BD205211.1 GI:33014981
KEYWORDS
JP 2002512813-A/1.
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1489)
Frechon,D.T.M., Laure,F.C. and Thierry,D.
Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL
Patent: JP 2002512813-A 1 08-MAY-2002;
BIORAD PASTEUR
OS Unidentified
PN JP 2002512813-A/1
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FH Key Location/Qualifiers
FT source 1..1489 /organism='Unidentified'.
FT
FEATURES             Location/Qualifiers
     source
     1..1489
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"

ORIGIN
Query Match          100.0%; Score 14; DB 6; Length 1489;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 397 GGCATCGTCAGTTG 410

RESULT 17
AX011297             1489 bp      DNA      linear      PAT 06-SEP-2000
LOCUS
DEFINITION
Sequence 1 from Patent WO9955908.
ACCESSION
AX011297
VERSION
AX011297.1 GI:9997847
KEYWORDS
Escherichia coli
SOURCE

```

```

ORGANISM             Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1
AUTHORS
Thierry,D., Frechon,D.T. and Laure,F.C.
TITLE
Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehec)
JOURNAL
Patent: WO 9955908-A 1 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
FRANCOISE CLAUDINE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES             Location/Qualifiers
     source
     1..1489
        /organism="Escherichia coli"
        /mol_type="unassigned DNA"
        /db_xref="taxon:562"

ORIGIN
Query Match          100.0%; Score 14; DB 6; Length 1489;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 397 GGCATCGTCAGTTG 410

RESULT 18
PSIS801/c            1517 bp      DNA      linear      BCT 07-JUL-2002
LOCUS
DEFINITION
P.syringae DNA for IS801 insertion sequence.
ACCESSION
X57269
VERSION
X57269.1 GI:45830
KEYWORDS
insertion element IS801.
SOURCE
Pseudomonas syringae
ORGANISM
Pseudomonadaceae; Pseudomonas.
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1517)
Romantschuk,M., Richter,G.Y., Mukhopadhyay,P. and Mills,D.
IS801, an insertion sequence element isolated from Pseudomonas
syringae pathovar phaseolicola
JOURNAL
Mol. Microbiol. 5 (3), 617-622 (1991)
MEDLINE
91260445
PUBMED
1646375
REFERENCE
2 (bases 1 to 1517)
Mills,D.
Direct Submission
AUTHORS
Submitted (23-JAN-1991) D. Mills, Oregon State University, Dept of
Botany and Plant Pathology, Corvallis OR 97331-2902, U S A
JOURNAL
Location/Qualifiers
FEATURES             Location/Qualifiers
     source
     1..1517
        /organism="Pseudomonas syringae"
        /mol_type="genomic DNA"
        /strain="pathovar phaseolicola, strain LR781"
        /db_xref="taxon:317"
     misc_feature
     1..5
        /note="insertion target sequence duplication"
        /evidence=experimental
     repeat_region
     6..1512
        /insertion_seq="IS801"
        101..1333
        /note="unnamed protein product; orf1"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA40540.1"
        /db_xref="GI:45831"
        /db_xref="GOA:P24607"
        /db_xref="UniProt/Swiss-Prot:P24607"
        /translation="MKPAYPLLQMSPAVTPRLKNLFTANQCAWHLLEGGLRDIE
VEVTKMLACGTSILGVKHYTCGNHSPHYKLYLCNTCHCRACPSGKKATQWITVON
NRLPDCPWQHLVFTLPDLPFLFFNRPWLQDALPRLAADNLIIYAKRGLRVGIFGAL
HTYGRRLNWHPHVHLSVTAGLDQGVWKNLSFHKALRRRMMWLVRDYLIGQPLSQL
TMPPLAHLICESDWRRLILAAGQGQWHIHLSSKTKNGRKTNYLGRYLLKPKPISGR

```

LAHYTNGATLRFTYLDHRTQAYQOETLSQADMLFRVVOHIEPKHFRMIRYFGLANRV
CGYLPRKYELKMATGPTPKLIFAPMAKAEFLNVDPRVCLCGARWVYTAISGLTV
387. .391
/note="putative"
complement(461. .1060)
/note="unnamed protein product; orf2"
/codon_start=1
/transl_table=11
/protein_id="CAA40541.1"
/db_xref="GI:45832"
/translation="MLPRDLVHLHKKVGLAQRLLIGLCAVIOVGEPORGPGVGMRO
TTARDGPPQVAHQVIGLSAVFRLLRQVDPVLATSGOYQTAQVATVEDVQRRHR
QLRKLSCQVIAHQHPQGLFVRQIFPDALLIQAHGRQVDVGMPEVPAIGV
QRPNPAQTAPGLGVQDVIGRQPKRIEQPAAVVEOQP"
misc_feature
1513. .1517
/note="insertion target sequence"
/evidence=experimental

ORIGIN

Query Match 100.0%; Score 14; DB 1; Length 1517;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 762 GGCATCGTCAGTTG 749

RESULT 19
AB022426 AB022426 2055 bp mRNA linear MAM 21-JAN-1999
LOCUS
DEFINITION Sus scrofa mRNA for FXII, complete cds.
ACCESSION AB022426
VERSION AB022426.1 GI:4165316
KEYWORDS FXII.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 2055)
Takahashi, T. and Kihara, T.
Porcine liver factor XII
Published Only in Database (1999)
REFERENCE
2 (bases 1 to 2055)
Takahashi, T. and Kihara, T.
Direct Submission
Submitted (07-JAN-1999) Takayuki Takahashi, Hokkaido University,
Graduate School of Science; Kitaku Kita 10 Joh Nishi 8 chome,
Sapporo, Hokkaido 060-0810, Japan
(E-mail:ttakahashi@sci.hokudai.ac.jp, Tel:81-11-706-2748,
Fax:81-11-706-2748)

FEATURES
source
Location/Qualifiers
1..2055
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue type="liver"
25. .1675
/codon_start=1
/product="FXII"
/protein_id="BAA37148.1"
/db_xref="GI:4165317"
/translation="MRALLLIGLVLSALLIPPKDKPRKKVMASEHTVLTVTG
EPCHFPPQYRLVYKICQGRGPRPWCATTPNFKDORWAYCLEPMKVDHCNKN
CQKGGTCVNMPPHICDPHFTGKCKQKCEFPQLFPQNEIWHREPAGVSK
CCKGKPAQCVASQVCNPNFLNGSCSLQTEGRLRCRPTGAGRLCDVDLKERY
SDRLGSLRWAGQYTLSPACQWASEATYWNMTAEALNWLGLGHAFRCPNDTRPM
CFVWRGQLNQWACQAPIGEAPPILPTQSPSEHQDSPILSREPQPTTPSQ
NLTSAWCAPPBPGRPLPSAGLVGGCGRKRLSLNRIVGLVLPQAHPIAALYWG
QNFCAGLIAPCVLTAAHCLONPAPPEELTVLIGDRHNQSCQCOTLAVRSYLHE
SYSPTQTQHDLLVRLKETADGCCAHPSPFPVQVCLPRSVASSAEPGALCEVAGWGH

QFCAEYSSFLQEAQVPLISPERCSAADVHGAAFTPGMLCAGFLEGTDACQDSSG
PLVCEDETAERQLVLRGIVSWGSGCGRDLKPGVYTDVANYLAWIQEHTTS"

ORIGIN

Query Match 100.0%; Score 14; DB 4; Length 2055;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 1771 GGCATCGTCAGTTG 1784

RESULT 20
AX073964/c AX073964 5145 bp DNA linear PAT 06-FEB-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0104295.
ACCESSION AX073964
VERSION AX073964.1 GI:12710224
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1
Avides, M.D., Deak, P. and Glover, D.M.
Orbit and homologues thereof
Patent: WO 0104295-A 1 18-JAN-2001;
University of Dundee (GB)
FEATURES
source
Location/Qualifiers
1..5145
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match 100.0%; Score 14; DB 6; Length 5145;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 3749 GGCATCGTCAGTTG 3736

RESULT 21
AF195498/c AF195498 5916 bp mRNA linear INV 01-NOV-2000
LOCUS
DEFINITION Drosophila melanogaster clone LD11488 Misexpression suppressor of
ras 7 (MESR7) mRNA, MESR7-3403 allele, complete cds.
ACCESSION AF195498
VERSION AF195498.1 GI:11066120
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 5916)
Huang, A.M., Rubin, G.M., Tsang, G., Evans-Holm, M. and Suh, C.
Full length Drosophila melanogaster cDNA sequence
Unpublished
2 (bases 1 to 5916)
Huang, A.M., Rubin, G.M., Tsang, G., Evans-Holm, M. and Suh, C.
Direct Submission
Submitted (18-OCT-1999) Molecular and Cell Biology, University of
California at Berkeley, 545 Life Sciences Addition Bldg., Berkeley,
CA 94720-3200, USA
FEATURES
source
Location/Qualifiers
1..5916
/organism="Drosophila melanogaster"
/mol_type="mRNA"

gene
CDS
/db xref="taxon:7227"
/chromosome="3"
/map="78C1-2"
/clone="LD11488"
1. .5916
/gene="MESR7"
/allele="3403"
771. .5245
/gene="MESR7"
/allele="3403"
/note="suppressor of activated Ras when overexpressed in the adult eye"
/codon start=1
/product="Wise expression suppressor of ras 7"
/protein_id="AAG28470.1"
/db xref="GI:11066131"
/translation="MAYRKPESLDLGFQIQMPKADMRVKVQLAEDLVFLSDDTNSIVC
TDMGFLIDGLMPWLTGSHFKIAQKSEAFSELIKRIGSDNFNATATVLPVHIDRLGDS
RDTVREKAQILLDLMEHRLVLPQALIDKLATSCFKHNAKRVREEFQITIVNALHEVG
TQQLSVRYVIPPVALLGDDTVNVRRAAIQTLVEIKYHGDRLRDLRMDVPAKLL
AMLEQFDQVQKGLLPALSALNTNGVGLDEADNIGLRERPTRMKRLPHLSAVSS
LRPKPNVDVTDGAVTMSFESSFEVVPQNLFIHAKOMDDIYKQVLVVISDKNADW
EKRVDAIKIRALLISYHTQPOFVAVQLKELSLFVDILKEELRSQVIREACITIA
MSKTRNLDAFCWSILEHLINLIQNSAKVIASTIALKYIIKYTHAPKLKIYDT
LNQSKDIRSTLCELMVLLEFQWOTKALERNATVLDTLKKSIGDADCDARHSRYA
YWAFFRHPFLAQIYGTLDIAAQRALEREREQGGGTGTGTAPETRTVSRIGR
TPGLQKPTPSMRSISAVDTAAQRAKRAQYTIYSQRKPLGPNNSNQASMTGAAS
GSLPRPLNSGGTTPATPGSVTPRPRAGVVSQPGSRSTSPSKLRDQYGGIGN
YRGATGAI PKASGIPRSTASRETSPTRSGGLMKRSMYSGAGSRRTPERNNPV
PSAAQLAQREAEHTLVGDDGQPDYVSGDYMRSGGMKRMGRKLMGRDSDIDSEA
SSVCSRSFSSYTRGNKSNYSLSGSHTRLDWSTORAPFDIETIQFCASTHWSERK
DGLISLTQYADGKELTQOQLKCVLDMFKMWDTHKVSLELDTVELLIVHANSEL
HEWLFIITLRLFNKGLTDLNLSMHSKIWKTLQVHYEFTQQLKELFRILISSTQTP
TTKTRIALRFLTDLANTYCKSSDFPSDQACERTVLKLAQAADQKSMELRSQARS
CLVALYNLANTPQMTLLADLPKVYQDSARSCIHSMMRQSCNCSGNSPSSPLSS
SPKLPQSPVGPFPASLQSHHQLSISSTSPRSQSSVEQELLFSSSELDIOHNIQKTSSE
EIRHCFGGQYQTALAPNGFNHQLSYHQDQGDQSCASLSSNKTQSSANTQSTNPESA
TMRLDLNERRTTQNAKSPDDAKVITVSINMAENGLIILASINMESEVVRVALLTK
DQVELLQTSLTNLGICIKGNCELPNKFPSIRMLNLILEAEHTDVIAGLHVLSK
IMRSNKRHNHMHFLELILIKIQCQHSKEALRDIDSMIPRIAPSLPLDLSINVP
VIATGEFPTNLCAIKILLEVTEHHGSEITDAHLDIVFPNLARSADDTQSMVRAAVFC
IVKLYFVLGEEKVKPKLSVLNPSKVLNLLNVYIEKQRNCISGGGSSTKNSAASS"

ORIGIN

Query Match 100.0%; Score 14; DB 3; Length 5916;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 4451 GGCATCGTCAGTTG 4438

RESULT 22
AF250842/c
LOCUS Drosophila melanogaster multiple asters (Mast) mRNA linear INV 04-AUG-2000
DEFINITION Drosophila melanogaster multiple asters (Mast) mRNA, complete cds.
ACCESSION AF250842
VERSION AF250842.1 GI:7650478
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5938)
REFERENCE Lemos,C.L., Sampaio,P., Maiato,H., Costa,M., Ome'l'yanchuk,L.V.,
AUTHORS Liberal,V. and Sunkel,C.E.
TITLE Mast, a conserved microtubule-associated protein required for
bipolar mitotic spindle organization
JOURNAL EMBO J. 19 (14), 3668-3682 (2000)
MEDLINE 20359266
PUBMED 10899121

REFERENCE 2 (bases 1 to 5938)
Maiato,H., Lemos,C.L., Sampaio,P. and Sunkel,C.E.
Direct Submission
Submitted (30-MAR-2000) Genetica Molecular da Mitose, Instituto de
Biologia Molecular e Celular, Rua Campo Alegre, 823, Porto
4150-180, Portugal
FEATURES
source
1. .5938
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="3"
/map="3L; 78C1-C2"
/clone="LD11488"
/dev stage="0-24 hours embryo"
1. .5938
/gene="Mast"
770. .5245
/gene="Mast"
/note="Mast; microtubule-associated protein; essential for centrosome separation and organization of mitotic apparatus"
/codon start=1
/product="multiple asters"
/protein_id="AAF66060.1"
/db xref="GI:7650479"
/translation="MAYRKPESLDLGFQIQMPKADMRVKVQLAEDLVFLSDDTNSIVC
TDMGFLIDGLMPWLTGSHFKIAQKSEAFSELIKRIGSDNFNATATVLPVHIDRLGDS
RDTVREKAQILLDLMEHRLVLPQALIDKLATSCFKHNAKRVREEFQITIVNALHEVG
TQQLSVRYVIPPVALLGDDTVNVRRAAIQTLVEIKYHGDRLRDLRMDVPAKLL
AMLEQFDQVQKGLLPALSALNTNGVGLDEADNIGLRERPTRMKRLPHLSAVSS
LRPKPNVDVTDGAVTMSFESSFEVVPQNLFIHAKOMDDIYKQVLVVISDKNADW
EKRVDAIKIRALLISYHTQPOFVAVQLKELSLFVDILKEELRSQVIREACITIA
MSKTRNLDAFCWSILEHLINLIQNSAKVIASTIALKYIIKYTHAPKLKIYDT
LNQSKDIRSTLCELMVLLEFQWOTKALERNATVLDTLKKSIGDADCDARHSRYA
YWAFFRHPFLAQIYGTLDIAAQRALEREREQGGGTGTGTAPETRTVSRIGR
TPGLQKPTPSMRSISAVDTAAQRAKRAQYTIYSQRKPLGPNNSNQASMTGAAS
GSLPRPLNSGGTTPATPGSVTPRPRAGVVSQPGSRSTSPSKLRDQYGGIGN
YRGATGAI PKASGIPRSTASRETSPTRSGGLMKRSMYSGAGSRRTPERNNPV
PSAAQLAQREAEHTLVGDDGQPDYVSGDYMRSGGMKRMGRKLMGRDSDIDSEA
SSVCSRSFSSYTRGNKSNYSLSGSHTRLDWSTORAPFDIETIQFCASTHWSERK
DGLISLTQYADGKELTQOQLKCVLDMFKMWDTHKVSLELDTVELLIVHANSEL
HEWLFIITLRLFNKGLTDLNLSMHSKIWKTLQVHYEFTQQLKELFRILISSTQTP
TTKTRIALRFLTDLANTYCKSSDFPSDQACERTVLKLAQAADQKSMELRSQARS
CLVALYNLANTPQMTLLADLPKVYQDSARSCIHSMMRQSCNCSGNSPSSPLSS
SPKLPQSPVGPFPASLQSHHQLSISSTSPRSQSSVEQELLFSSSELDIOHNIQKTSSE
EIRHCFGGQYQTALAPNGFNHQLSYHQDQGDQSCASLSSNKTQSSANTQSTNPESA
TMRLDLNERRTTQNAKSPDDAKVITVSINMAENGLIILASINMESEVVRVALLTK
DQVELLQTSLTNLGICIKGNCELPNKFPSIRMLNLILEAEHTDVIAGLHVLSK
IMRSNKRHNHMHFLELILIKIQCQHSKEALRDIDSMIPRIAPSLPLDLSINVP
VIATGEFPTNLCAIKILLEVTEHHGSEITDAHLDIVFPNLARSADDTQSMVRAAVFC
IVKLYFVLGEEKVKPKLSVLNPSKVLNLLNVYIEKQRNCISGGGSSTKNSAASS"

ORIGIN

Query Match 100.0%; Score 14; DB 3; Length 5938;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTTG 14
| | | | | | | | | | | | | | | |
Db 4450 GGCATCGTCAGTTG 4437

RESULT 23
AB031048/c
LOCUS Drosophila melanogaster orbit mRNA for microtubule
DEFINITION associated-protein orbit, complete cds.
ACCESSION AB031048
VERSION AB031048.1 GI:7527325
KEYWORDS Microtubule associated-protein orbit.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

SPKLPSPVGPASLQSHHQLSISSTSPRSQSSVQEQLLPSSBLDIQHNQIKTSE
EIRHCFGQYQVOTALPANGFNHGLQYHQDQDSCASLSNSKTQSSANTTQSNTPESA
TWRLDLNERITQNAKSPDDAKVITVSINMAENGELILASMESEVVRVALLITLK
DQVELLQTSLTNIGICIKGNCDELPNKFRSIRMLNLILEAEDTDVVIAGLVLISK
IWSNKRNNHMFLEILLIKITQCYQHSKEALRDSIMIPRIAPSLDLSINIVNP
VIATGFPNLCALKILLEVTEHHGSEITDAHLDIVFPNLARADDTQSMVRAAVFC
IVKLYFVLGEEKVKPLSVLNPSPKVRLLNVIYIEKQNCISGGGSSITKNSSAASSS"

ORIGIN

Query Match 100.0%; Score 14; DB 3; Length 5962;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTTG 14
|||||
Db 4446 GGCATCGTCAGTTG 4433

RESULT 25
AE008914/c
LOCUS
DEFINITION
Salmonella typhimurium LT2, section 218 of 220 of the complete genome.
ACCESSION
AE008914 AE006468
VERSION
AE008914.1 GI:16423108
KEYWORDS
SOURCE
ORGANISM
Salmonella typhimurium LT2
Salmonella typhimurium LT2
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.

REFERENCE
1 (bases 1 to 22694)
McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W.,
Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F.,
Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A.,
Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W.,
Stonking, T., Nhan, M., Waterston, R. and Wilson, R.K.
Complete genome sequence of Salmonella enterica serovar Typhimurium LT2

JOURNAL
Nature 413 (6858), 852-856 (2001)
MEDLINE
21534948
PUBMED
11677609
REFERENCE
2 (bases 1 to 22694)

The Salmonella typhimurium Genome Sequencing Project
Direct Submission
Submitted (29-MAR-2001) Genome Sequencing Center, Department of
Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT
Supported by NIH grant 5U 01 AI433283

Coding sequences below are predicted from manually evaluated
computer analysis, using similarity information and the programs;
GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and
GeneMark; <http://opal.biology.gatech.edu/GeneMark/>
EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>,
and Pedro Romero and Peter Karp at EcoCyc;
<http://ecocyc.PangeaSyste.ms.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites
were kindly provided by Heladia Salgado, Julio Collado-Vides and
ReguondB;
http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one ml3 subclone.
Location/Qualifiers

FEATURES
source
1. .22694

/organism="Salmonella typhimurium LT2"
/mol_type="genomic DNA"
/strain="LT2; SGSC 1412; ATCC 700720"
/db_xref="ATCC:700720"
/db_xref="taxon:99287"
/notes="LT2"
complement(98. .2350)
/gene="mdoB"
/note="synonym: STM4541"
complement(98. .2350)
/gene="mdoB"
/EC number="2.7.8.20"
/note="similar to E. coli phosphoglycerol transferase I (AAC77315.1); Blastp hit to AAC77315.1 (750 aa), 92% identity in aa 1 - 749"
/codon_start=1
/transl_table=11
/product="phosphoglycerol transferase I"
/protein_id="AAL23359.1"
/db_xref="GI:16423109"
/translation="MLIYAKAGRTMFAATLTVLGLFVILNITLYASDYFTGGIN
DAVYLTNLSLGTAGVKYILPGIGIALALVAVFGALGWLRHRRHHPHVGYSLLAL
LLALGSDASPAFPQITELVQSMDGDDFAVYKPEAKTIIPKLNLIYIVGESLE
RTYFDNAPFNLTPELGALKNGLDFTHTQGLDFTYTAGMVASQCGIFLFPFEGN
ASASVSPFPNITCLGDIILKNISQYNYFVQGANLRFAGKDVFLKSHGFIDHLYGAEELK
TVADPSYRNDWGYDDTVLDEAWKFEALSRSGQSFSLFTLVTDTHDPDFISRTCN
RKVYDYGKPNOSFSAVCSQENIAEFINKIKASPMFKDTVIIVSSDHLAMNNTAWKY
LNKQDNNLFFILRGKPKQETLAVKNTMDNGATVLDILGGDNFTGLGRSSLSGQSL
SEVFLNKKVLAAMPDIIIRLMNFPKDEIDTVDKDNMIAFGSGHFRLLLRVSDK
RVPELESYSPALFQDLADFAFDNFVWIDRCYKQALWAPALALSTDMCVSQGQK
GOOTVQHVDAKQWQGTAKFDKTMIDERYKGNVDITLKVNDIRYKADSFIFNVQAGP
EEVKQFSGISRPSEWGRWNAQLGDEVKIEYKAPLKKFLDVTAKAFGNSANRPV
RVGNEQTLVLGHVDVTIILHFNPTDANTLVIAPPVSTNEGNIILGHSPKRLIGIM
VEIKVNVES"
complement(2672. .3145)
/gene="YjJ"
/note="synonym: STM4542"
complement(2672. .3133)
/gene="YjJ"
/note="similar to E. coli putative glycoprotein/receptor (AAC77316.1); Blastp hit to AAC77316.1 (165 aa), 68% identity in aa 3 - 165"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/protein_id="AAL23360.1"
/db_xref="GI:16423110"
/translation="MNTAKHVLCCAAIASVLISTGIAASQLSAQAGANAQCGMSLSA
LTGLLSRGASQLSADNMNNAAGILQYCAKOKLASATNVENVKQILNKLGLDITQQSQ
DTNVLNGQLLTKDKGQQLNANNIGSTPLAEKVKTKACDLVLQQLGLNFLS"
complement(3140. .3145)
/gene="YjJ"
/note="putative RBS for YjJ; RegulonDB:STWS1H004408"
complement(3191. .3947)
/gene="dnaC"
/note="synonym: STM4543"
complement(3191. .3928)
/gene="dnaC"
/note="initiation and chain elongation; similar to E. coli chromosome replication; initiation and chain elongation (AAC77317.1); Blastp hit to AAC77317.1 (245 aa), 93% identity in aa 1 - 245"
/codon_start=1
/transl_table=11
/product="chromosome replication protein"
/protein_id="AAL23361.1"
/db_xref="GI:16423111"
/translation="MKNVGLMQLRQKMPAHITPAFKTGBELIAWQEGEIRAAL
ARNRANKMORTNRSGIRPLHONCSFDNVRECDGMNALSKARQVDFPDGNIASF
VFGKCPGTGKHLAAATCNELLARGSVLIITVADINSAMKDTFNSRETSEQLNDL
SNVDLLVIDEIGVQTSRYEKVLIQIVDRSSSRKRTGMLTNSNMEETMQLSERVM
DMRLGNLSLWNFTWDSYRSRVTKKEY"
complement(3931. .4480)

gene


```

CDS
/ gene="dnaT"
/ note="synonym: STM4544"
/ complement(3931..4470)
/ gene="dnaR"
/ note="similar to E. coli DNA biosynthesis; primosomal
protein i (AAC77318.1); Blastp hit to AAC77318.1 (179 aa),
81% identity in aa 1 - 179"
/ codon_start=1
/ transl_table=11
/ product="primosomal protein I"
/ protein_id="AAL23362.1"
/ db_xref="GI:16423112"
/ translations="MSSRIILTSDVIGDVLHDDHVAIAKSTGCAVAFVANNAPAFVA
VTPARMAELLALSEKISPPGSDVALDAQFYEEPEAAAPVAIPCGKFAFMYPAWQPDADPQ
RQAALGVALREPVTAEEAFAFYAQEGKVFFHHIQWQKRLAKSVQISRSSNGGMPQ
RDINSVSEPDNHHIPPPRG"
/ complement(3942..3947)
/ gene="dnaT"
/ note="putative RBS for dnaC; RegulonDB:STWS1H004409"
/ complement(4475..4480)
/ gene="dnaI"
/ note="putative RBS for dnaT; RegulonDB:STWS1H004410"
/ complement(4563..5051)
/ gene="STM4545"
/ complement(4563..5036)
/ gene="STM4545"
/ notes="similar to E. coli orf, hypothetical protein
(AAC77319.1); Blastp hit to AAC77319.1 (108 aa), 89%
identity in aa 1 - 108"
/ codon_start=1
/ transl_table=11
/ product="putative inner membrane protein"
/ protein_id="AAL23363.1"
/ db_xref="GI:16423113"
/ translations="MGIIDFLALMQDMILSAIPAGFAMFVFNPHRALPWCALLGAL
GHGSRMLMWSAGFNIESTFMSALLVGSIGIOWSRWYLAHPKVTFVAIVPMPPGISA
YTAMISAVKISHLGYSEPPMITLLTNFLKASSIVGALSIGLSVPLGLWLYKRPV"
/ complement(5027..5953)
/ gene="yjjp"
/ note="synonym: STM4546"
/ complement(5027..5938)
/ gene="yjjp"
/ notes="similar to E. coli putative structural protein
(AAC77320.1); Blastp hit to AAC77320.1 (277 aa), 90%
identity in aa 19 - 277"
/ codon_start=1
/ transl_table=11
/ product="putative membrane protein"
/ protein_id="AAL23364.1"
/ db_xref="GI:16423114"
/ translations="MRLAVDVFINKKILTNAMPVLRVKKEASVIVAHRMSFRTRG
VMQEQSIQRAVRLCIQCGLFLLQGAESALVDELSTRGLAGMDSVESAISNAI
VLTTIKDGOCLSTRKNQDRGINHHVVEVQHIVILAEHRLDYKGVEKRFSQLRPLR
YRWLVAFMYGLSCFCCKNNGWDGAVITFFASMIAMYIROMLAQRHLHPQINFCCI
TAPVATISGLMTLPRAFSTQPTIAMAASVLLLVPGFPLINSVADMEKGHINTGLARW
AIASLLTLCVGVMSMTVWGLRGV"
/ complement(5046..5051)
/ gene="STM4545"
/ note="putative RBS for STM4545; RegulonDB:STWS1H004411"
/ complement(5948..5953)
/ gene="yjjp"
/ note="putative RBS for yjjp; RegulonDB:STWS1H004412"
/ complement(5908..7241)
/ gene="yjjp"
/ note="synonym: STM4547"
/ complement(6508..6513)
/ gene="yjjp"
/ note="putative RBS for yjjp; RegulonDB:STWS1H004413"
/ complement(6516..7241)
/ gene="yjjp"
/ notes="similar to E. coli putative regulator (AAC77321.1);
Blastp hit to AAC77321.1 (241 aa), 66% identity in aa 1 -
241"

```

```

/codon_start=1

Query Match      100.0%; Score 14; DB 1; Length 22694;
Best Local Similarity 100.0%; Pred. NO. 1.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGCATCGTCAGTTG 14
      |||||
Db      20877 GGCATCGTCAGTTG 20864

Search completed: March 11, 2005, 06:55:16
Job time : 81.4641 secs

```

This Page Blank (uspto)

Mon Mar 14 11:04:14 2005

/organism='Unidentified'.

```

FEATURES
  source
    1..1489
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"

ORIGIN
  Query Match 100.0%; Score 31; DB 6; Length 1489;
  Best Local Similarity 100.0%; Pred. No. 0.0081;
  Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
  |||
  1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
Db 387 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 417

RESULT 4
AX011297 1489 bp DNA linear PAT 06-SEP-2000
LOCUS
DEFINITION Sequence 1 from Patent WO955908.
ACCESSION AX011297
VERSION AX011297.1 GI:9997847
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1
AUTHORS Thierry,D., Frechon,D.T. and Laure,P.C.
TITLE Nucleotide sequences for detecting enterohemorrhagic escherichia coli (ehc)
JOURNAL Patent: WO 9955908-A 1 04-NOV-1999;
THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

FEATURES
  source
    1..31
    /organism="Escherichia coli"
    /mol_type="unassigned DNA"
    /db_xref="taxon:562"

ORIGIN
  Query Match 100.0%; Score 31; DB 6; Length 31;
  Best Local Similarity 100.0%; Pred. No. 0.0074;
  Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
  |||
  1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
Db 1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31

RESULT 3
BD205211 1489 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION BD205211
VERSION BD205211.1 GI:33014981
KEYWORDS JP 2002512813-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1489)
AUTHORS Frechon,D.T.M., Laure,P.C. and Thierry,D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 1 08-MAY-2002;
BIORAD PASTEUR
OS Unidentified
PN JP 2002512813-A/1
PD 08-MAY-2002
PF 27-APR-1999 JP 2000546051
PI 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08 C07K14/245, C12N1/21, C12N15/09, C12Q1/68, C12N15/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
(EHEC).
CC Key Location/Qualifiers
FH Key 1..1489
FT source

```


<p>TITLE Direct Submission</p> <p>JOURNAL Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA</p> <p>FEATURES Location/Qualifiers</p> <p>source</p> <p>1..93077</p> <p>/organism="Escherichia coli O157:H7"</p> <p>/mol_type="genomic DNA"</p> <p>/strain="EDL933"</p> <p>/serotype="O157:H7"</p> <p>/lab_host="Escherichia coli C600"</p> <p>/plasmid="pO157"</p> <p>1..561</p> <p>/gene="fno"</p> <p>/notes="synonym: L7001"</p> <p>1..561</p> <p>/gene="fno"</p> <p>/notes="97 pct identical amino acid sequence and equal length to F101_ECOLI SW: P22707"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="fertility inhibition protein (conjugal transfer repressor)"</p> <p>/protein_id="AAC70069.1"</p> <p>/db_xref="GI:3822115"</p> <p>/translation="MAEQKRPVLTLLKRTGETPVRSRKTIINVTTPKWKVKOKLA EKAARAEALAAQAQALSIYLNLPQDEAVNTLKPWPGFDGDTPLLAGGIRD VLEEDVAQRNIPLSHKKLRALKAIRSESYLCAMKAGACRYTEGYVTEHISQEEA YAAARDLKDINQRKIKAEIQAVLDEK"</p>	<p>promoter</p> <p>653..681</p> <p>/notes="predicted sigma 70 promoter; score of 52%"</p> <p>698..949</p> <p>/gene="L7002"</p> <p>698..949</p> <p>/gene="L7002"</p> <p>/notes="54 pct identical (1 gap) to 23 residues of an approx. 269 aa protein EXOK_RHIME SW: P33693"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="unknown"</p> <p>/protein_id="AAC70070.1"</p> <p>/db_xref="GI:3822116"</p> <p>/translation="MDSETVHGTVRSQVTSVPEAGPLFWKSVDAWKQKRGDGLPVL HPGLTGSSLPKGLNTATGAERGGKESLSHYRDSRG"</p> <p>1025..1052</p> <p>/note="predicted sigma 70 promoter; score of 62%"</p> <p>1151..1612</p> <p>/gene="L7003"</p> <p>1151..1612</p> <p>/gene="L7003"</p> <p>/notes="98 pct identical and equal length to YF13_ECOLI SW: Q99342"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="hypothetical protein 15.6 kDa protein in fno 3' region precursor"</p> <p>/protein_id="AAC70071.1"</p> <p>/db_xref="GI:3822117"</p> <p>/translation="MRKYIPLVLFISWPVLCADHGRVVRVLDGDTIEVMSRKAIV IRLVNDAPKKQDYGKRWSTDMKSLVAGTKVTVTYFORDRYGRMLGOVYAPDGMNVN QFMVRAGAAWVYEQYNTDPLVPLQNEARQKGLMSDADVPVPMIWRHRK"</p> <p>1390..1419</p> <p>/note="predicted sigma 70 promoter; score of 56%"</p> <p>1658..1867</p> <p>/gene="L7004"</p> <p>1658..1867</p> <p>/gene="L7004"</p> <p>/notes="55 pct identical (0 gaps) to 66 residues of an approx. 72 aa protein HHA_ECOLI SW: P23870"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="putative hemolysin expression modulating protein"</p>	<p>promoter</p> <p>1899..1927</p> <p>/note="predicted sigma 70 promoter; score of 56%"</p> <p>1905..2243</p> <p>/gene="L7005"</p> <p>1905..2243</p> <p>/gene="L7005"</p> <p>/note="98 pct identical (0 gaps) to 57 residues of an approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF 4, TRSW: Q52340"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="hypothetical protein"</p> <p>/protein_id="AAC70073.1"</p> <p>/db_xref="GI:3822119"</p> <p>/translation="MKLIIFILVILIAALLIRILSRVNOHSPLLMQLHAAGIRTDG ABRILSGGYDASHLRPERRRRDILLEVLQRQAGIPLLRSHDARKLLQMTGWLTTG AAQSPHERS"</p> <p>complement(2240..2327)</p> <p>/note="initiation site; Escherichia coli plasmid R100 ssIB gene; 99 pct identical to ssIB locus R10SSIB accession D90185"</p> <p>2416..2421</p> <p>/note="100 pct identical (0 gaps) to the -35 region at 23..28 locus ECNR1REP accession X02302"</p> <p>2439..2445</p> <p>/note="100 pct identical (0 gaps) to the -10 region at 46..52 locus ECNR1REP accession X02302"</p> <p>2473..2478</p> <p>/note="100 pct identical (0 gaps) to RBS at 80..85 locus ECNR1REP accession X02302"</p> <p>2483..2737</p> <p>/gene="cpb2"</p> <p>/note="synonym: L7006"</p> <p>2483..2737</p> <p>/gene="cpb2"</p> <p>/note="98 pct identical amino acid sequence and equal length to CPB2_ECOLI SW: P03847"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="CopB protein (RepA2 protein)"</p> <p>/protein_id="AAC70074.1"</p> <p>/db_xref="GI:3822120"</p> <p>/translation="MSQENAVTSSGAKRAYRKGNPLSDAEKQRLSVARKASPKREV KVFLEPKYKAMLMQMCHEDGLTQAEVLTAIKSEAQRVCV"</p> <p>2754..2782</p> <p>/note="predicted sigma 70 promoter; score of 69%"</p> <p>2973..3047</p> <p>/gene="L7007"</p> <p>2973..3047</p> <p>/gene="L7007"</p> <p>/note="91 pct identical (0 gaps) to 24 residues of an approx. 24 aa protein uORF_P30REPFC, accession M16167, translationally coupled to replication initiation protein"</p> <p>/codon_start=1</p> <p>/transl_table=11</p> <p>/product="replication initiation protein"</p> <p>/protein_id="AAC70075.1"</p> <p>/db_xref="GI:3822121"</p> <p>/translation="MLGKVQDFFLCGLLRIVSAGWCD"</p> <p>3032..3035</p> <p>/note="100 pct identical (0 gaps) to RBS at 641..644 locus ECNR1REP accession X02302"</p> <p>3040..3897</p> <p>/gene="rep2"</p> <p>/note="synonym: L7008"</p> <p>3040..3897</p> <p>/gene="rep2"</p> <p>/note="96 pct identical amino acid sequence and equal length to REP2_ECOLI SW: P03066"</p>
---	---	--

This Page Blank (uspto)

```

/codon_start=1
/transl_table=1
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/translation="MTDLOQTYRQVKNPNVFTFREGAGTLKFCCKMEKAVGFTSR
FDFAIYAHARSGLRRMPVLRRAIDALLOGLCFHYDPFANRVQCSITTLAIECG
LATSAAKLSITRALATFLSELGITVYQTEYDPLIGCVIPDTITFTALFAALDV
SEDAVAARSRRVEMENLRKKOGLDGLMDELAKAWRVFRERFRSYQTELKSRGFK
RARARDAGRRQDITLVKRLTREISGRFTANREAVKREVERRVKERMILSRNEN
YSRLATASP"
misc_feature
3847..4095
/standard_name="CIS"
/note="88 pct identical to locus ECCIS accession X12587,
required for cis-activation of oriR by the replication
initiation protein"
4072..4080
/note="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1RPP at (1682..1690) accession X02302"
4084..4232
/standard_name="oriP"
/note="89 pct identical to oriR (1094..1242); minimum
segment for replication of E. coli incFII plasmid NR1
ECREPA1 X12776"
/direction=right
4169..4197
/note="predicted sigma 70 promoter; score of 56%"
4259..4453
/gene="L7009"
4259..4453
/gene="L7009"
/note="95 pct identical to (0 gaps) 64 residues of a 128
aa protein REPA4 locus ECRS1 accession V00351"
/codon_start=1

Query Match      100.0%; Score 31; DB 1; Length 92077;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
|||||
DB 7223 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 7253
|||||

RESULT 6
AX191727 92077 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 9 from Patent WO0149775.
ACCESSION AX191727
VERSION AX191727.1 GI:15209896
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 Iversen, P.L.
AUTHORS
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 9 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
Location/Qualifiers
source
1..92077
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match      100.0%; Score 31; DB 6; Length 92077;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
|||||

```

Db 7223 AGCACTCAACGGCATCGTCAGTTGCGGCTTG 7253

```

RESULT 7
AB011549 92721 bp DNA circular BCT 27-APR-1999
LOCUS
DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION AB011549
VERSION AB011549.2 GI:4589740
KEYWORDS ToxR-regulated lipoprotein; tagA.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (sites)
AUTHORS Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K.,
Yutsudo, H. C., Kubota, Y., Yamauchi, Y., Iida, T., Yamamoto, K.,
Honda, T., Han, C. G., Ohtsubo, E., Kasamatsu, M., Hayashi, T., Kuhara, S.
and Shinagawa, H.
TITLE Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
outbreak
JOURNAL DNA Res. 5 (1), 1-9 (1998)
MEDLINE 98290540
PUBMED 9628576
REFERENCE 2 (bases 1 to 92721)
AUTHORS Makino, K.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1998) Kozo Makino, Research Institute for
Microbial Diseases, Osaka University, Molecular Microbiology,
Yamadaoka, 3-1, Suita, Osaka 562, Japan
(E-mail:makino@biken01.biken.osaka-u.ac.jp, Tel:81-6-879-8318,
Fax:81-6-879-8320)
COMMENT On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES
Location/Qualifiers
source
1..92721
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:562"
/plasmid="pO157"
/note="RIMD 0509952 is a strain of enterohemorrhagic E.
coli, EHEC O157:H7"
join(92527..92721,1..2502)
/gene="tagA"
join(92527..92721,1..2502)
/gene="tagA"
/codon_start=1
/transl_table=1
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4666293"
/translation="MTKMERWETPMKLYLSCTILAPLAIGVFSATAADNNSAIYF
NTSQINDLOSLLAEVYKPAQSLPAHPREGDSQPHLTSRKSLLLVLPVKADDKTP
VQVEARDNNKILGTLTYIPSSLPDTIYHLDGVPEGIDTFPHNGTKKIIINTVAEVN
KLSADSGSIHSHLTNNALVEIHTANGRWVRDIYLPQPDLEGKMRVFSVAGYSSTV
FYGRKVTLSVGTLLFKYVNGQWFRSGELNNRITYAQHIWSAELPAHWIPLVGLNV
IKQNLGRLLNDIKIGAPGELLHTIDIGMLTTPRDFPAKDEAHREYFTIPIVSR
MTVNYAPLHKVYMLPTGELLTMDPQNGMHSQGTMRQIKGLVSHGIDNANYGLN
STAGLGENSHPIYVVAQAANNSRGVANGIQVHSGSGGGIYTLDSLTGNEFSHEVGH
NYGLGHYVDFPGKSVHSAENNNNTWGDGKKRPIPNFYPSQTSKSCLNQCQEPH
DGHKFYDGAAGGSPFSAANRFTWYTPNSSAIIOQFPENKAVFDSRSTGSKMNADT
QEMEPYBHTIDAEQITASVNELSESKMAELMAEYAVVKVMWNGNWTNNIYIPTASA
DNRSILTIINHEAGNSYLFINGDEKVSQYKKSFSVSDQGFVSDQFVDRVYDTRAKPE
QFGVPTTLVGYDPEGLTSSYIYPMAYGYFTSYSDSNDLSNDCQLQVDYTKEGQL
RPLRANRANNTVANKHINYPTEOPTQATLVNCKNLTDKSLTPAEGLYTVNGQ
ALPARENEGCIVSVNSGKRYCLPVQSGSGSLPDIWGVQEVYVDSGAKAKVLLSDMDN
LSYNRIGFVGNVNPADMKVKAWNGQYLDPSKPSMRVYTK"
2589..3464
/gene="etpC"
2589..3464
/gene="etpC"

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 19703.4 Seconds
(without alignments)
2876.537 Million cell updates/sec

Title: US-09-674-277-1

Perfect score: 1489

Sequence: 1 ctgcagccgagatgaag.....ctggaaggagcctggtcgac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_est1.*

2: gb_est2.*

3: gb_hcc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_g881.*

9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	227.4	15.3	717	8	BZ893488 HL3_0183
2	201	13.5	757	6	CD425096 SAL_10_G0
3	200	13.4	701	6	CD422961 SAL_38_H1
4	197.2	13.2	762	5	BQ752115 ESTG32678
5	196	13.2	707	5	BQ751801 ESTG32364
6	191	12.8	707	5	BQ751801 ESTG32364
7	190	12.8	600	8	BZ893918 HL8_0153
C 8	185.2	12.4	742	5	BQ751881 ESTG32444
9	183.2	12.3	706	5	BQ751387 ESTG31950
10	183	12.3	726	7	CF876888 trico074xa
11	183	12.3	780	6	CB905354 trico074xa
12	182.4	12.2	721	7	CF869061 trico017xj
13	182.4	12.2	726	7	CF870162 trico021x1
14	182.4	12.2	748	7	CF869494 trico019x1
15	182.4	12.2	776	6	CB899177 trico017xj
16	182.4	12.2	790	6	CB900330 trico021x1
17	182.4	12.2	803	6	CB899636 trico019x1
18	181	12.2	827	6	CB901454 trico026xk
19	181	12.2	827	7	CF871230 trico026xk
20	179.2	12.0	598	6	CD426092 SAL_16_C0
21	179.2	12.0	620	7	CN476380 SPVD16A09
22	178.6	12.0	741	6	CD426471 SAL_20_F1
C 23	177.4	11.9	720	8	BZ892154 Hml_0140.
24	176.8	11.9	729	2	AW179968 MGA0020f

25	176.4	11.8	562	4	BG370112	BM370112 Lma123 V8
26	171.8	11.5	737	9	CNS041HH	AL14438 Anopheles
27	171.4	11.5	751	7	CF869797	trico020xh
28	171.4	11.5	807	6	CB899943	CB899943 trico020xh
29	166.4	11.2	628	1	AJ636328	AJ636328 trico020xh
30	162	10.9	826	7	C0064984	est_k_bre
31	161	10.8	573	2	AW180683	AW180683 MGA031f
32	159.4	10.7	465	4	BG278355	BG278355 a390np.r
33	158.6	10.7	805	7	C0064223	est_k_bre
34	158.4	10.6	663	4	BM865076	BM865076 mgap008xD
35	154.6	10.4	546	5	BQ142683	BQ142683 113 Metar
C 36	152.6	10.2	786	8	BH403595	BH403595 AG-ND-127
37	140.2	9.4	772	7	C0060620	est_k_bre
38	134.4	9.0	544	4	BM869928	BM869928 mgns008xE
39	134.2	9.0	571	2	AW180185	AW180185 MGA0268E
40	132.6	8.9	600	2	AW180270	AW180270 MGA0361F
41	131.8	8.9	738	7	CF877907	CF877907 trico077xk
42	131.8	8.9	794	6	CB906103	CB906103 trico077xk
43	131.4	8.8	668	7	C0059563	est_k_bre
44	130.4	8.8	703	7	C0063988	C0063988 est_k_bre
45	124.8	8.4	720	7	CF867685	CF867685 trico012xd

ALIGNMENTS

RESULT 1
BZ893488/c
LOCUS BZ893488 717 bp DNA linear GSS 30-JUL-2003
DEFINITION HL3_0183 H1 pUC18 Library Halorubrum lacusprofundi genomic 5',
genomic survey sequence.
ACCESSION BZ893488
VERSION BZ893488.1 GI:33344078
KEYWORDS GSS.
SOURCE Halorubrum lacusprofundi
ORGANISM Halorubrum lacusprofundi
Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
Halobacteriaceae; Halorubrum.
REFERENCE 1 (bases 1 to 717)
AUTHORS Goo,Y., Roach,J., Glusman,G., Baliga,N.S., Deutsch,K., Pan,M.,
Dasgupta,S., Ng,W.V. and Hood,L.
TITLE Low-pass Sequencing for Microbial Comparative Genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Goo Y
Institute for Systems Biology
1441 North 34th Street, Seattle, WA 98103, USA
Tel: 206 732 1412
Fax: 206 732 1299
Email: ygoosystemsbiology.org
Seq primer: M13 Forward
Class: Shotgun.

FEATURES

source
1. 717
/organism="Halorubrum lacusprofundi"
/mol_type="genomic DNA"
/strain="ATCC 49239"
/db_xref="taxon:2247"
/clone_lib="H1 pUC18 Library"
/note="Vector: pUC18; Site 1: SmaI; A shotgun library was
constructed from Halorubrum lacusprofundi genomic DNA
using pUC18/SmaI/BAP plasmid"

ORIGIN

Query Match 15.3%; Score 227.4; DB 8; Length 717;
Best Local Similarity 61.5%; Pred. No. 2.9e-58;
Matches 404; Conservative 0; Mismatches 241; Indels 12; Gaps 2;
Qy 723 TTTCACAGCTGGATATGAGGCTCTGAAAAAGATATCAAGATTTCTGACACTTCC 782
Db 678 TTTCAGGAGCTCGACCTCGATCGGTGAAAGCGATATCGAGGACGTAATGACGACATCG 619
Qy 783 CAGGATGGTCCCTCGGATATGGTCATATGCTCTTCTTATTCGATCGCTGG 842

618 CAGGACTGGTGGCGCGGACTACGACACCTACGCGCGCTTTTCATCCGATGGCGTGG 559
Qy 843 CACGGTCCGGAAACATACAGGACATATGATGGCCGGGAGGCGCCAGTGGTGCACAA 902
Db 558 CACAGCCCGGTACGTACCGGACGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 499
Qy 903 CGTTTGAACCGCTGACGCTGCGCGGATTAACCTTAATCTGGATAAAGCCGTCGATTG 962
Db 498 CGGTCCACCGCTTAACAGCTGGCGCGACACCTCAACCTTCGACAGCGCGCGGACTG 439
Qy 963 CTGTGGCAGTCAAGAAAAATACGGCTCCAGTATTTCTGGGAGACCTGATGGTCTG 1022
Db 438 CTCTGGCGGTCAACAGAGTACGGTCGCAAGCTCTCGTGGCGGATCTAATGTTCTC 379
Qy 1023 ACTGGTAATGTTCCTTGAATCATCGGATTTAAACCGTGGGATTTCTGGCGGAGA 1082
Db 378 ACTGGGAACGTGCGCTCGAGTCGATGGCTTCGAGACGTTTCGGCTTCGCGCGCGCG 319
Qy 1083 GAAGATCACTGGGAGTC---GGACTGTGTACTGCGGGGCTCGACAA---G 1130
Db 318 GAGGACAGGTTCAAAGCCGACGCGCTGACCTGGGGCCCCGAGAGCAATGGGAGTCG 259
Qy 1131 CCTCTGCAGATAACCGGGATAAAAACGGGAAACCTTCAGAAACCTTTGCCCGCCACGCG 1190
Db 258 AGCTCGCGGAGCGTTTCGACGAGGGAGGAGCTTCGACGAGGAACTCGGCAACACCGTC 199
Qy 1191 ATGGGACTTATTTATGTAATCTGAAGGCGCGCGGTGGAACACAGATCTCTCGCTTCC 1250
Db 198 ATGGGCTCTCTACTGTAATCCGAGGCGCGGAAACCGCGAGCGCGATCTGGAAGGTCC 139
Qy 1251 GCGAAGATATCAGGAGGAGCTTTTTCAGTATGCCATGGATGATGAGGAGCTGTGCC 1310
Db 138 GCGGCGAATATCCCGACACGCTTCAGCCACATGCGGATGAACAGCAAGGAGCGTCCGA 79
Qy 1311 CTGATCGCGGAGGCGCATACATTTGGTAAAGCACATGCTGAGCGCTCTCTGAAAAA 1367
Db 78 CTATCGCGCGGTTCACACCTTCGGCAAGTTCACGGCGCGGACTCGGCGGACAA 22

CD425096 757 bp mRNA linear EST 02-JUN-2003
LOCUS SAI_10_G08.gi_A002 Salicylic acid-treated seedlings Sorghum bicolor
DEFINITION cDNA clone SAI_10_G08_A002 5', mRNA sequence.
ACCESSION CD425096
VERSION CD425096.1 GI:31331359
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 757)
Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S.,
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Salzman, R.,
Chua Tan, N., Gonzalez, M., Lane, S., Miller, V., Nanda, P.,
Olaseinde, O., Eastman, A., and Pratt, L.H.
An EST database from Sorghum: salicylic acid-treated seedlings
Unpublished (2003)
Other ESTs: SAI_10_G08.b1_A002
Contact: Cordonnier-Pratt, MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics;
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences

are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAAAGCTGCG).
FEATURES
source
1. 757
/organism="Sorghum bicolor"
/lab host="DHI08-T1 phage-resistant E. coli"
/clone lib="Salicylic acid-treated seedlings"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. At 8 days of age, medium was
supplemented with 1 mM salicylic acid (SA). Roots and
shoots were harvested after 27 and 72 hr and material from
both time points was combined prior to RNA isolation.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5'-prime
DraIII site is CACTGTGTG, 3'-prime DraIII site is
CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 13.5%; Score 201; DB 6; Length 757;
Best Local Similarity 67.1%; Pred. No. 4.4e-50;
Matches 285; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 678 CTTGAATCAAATCCTCGGGGGCTGATTTGATATGTCACAGATTTCAACAGCTGGAT 737
Db 220 CCTCAGTATAACCATTTGGGTGCTGATTCGACTACACCGAGGCTTTCAAGTCTCTGGAC 279
Qy 738 ATGGAGGCTCGAAAAAGATATCAAGATTTGCTGACNACTTCCAGGATTTGGTCCCT 797
Db 280 TTTGCTCATTTGAAGAAGGACCTCAACGCGCTTCTCACTGATTTCTCAGGATTTGGTGGCT 339
Qy 798 GCGGATTTATGCTATTATGCTCTTTCTTTATTTGCTATGGCTTGGCGAGTGGCGGAACA 857
Db 340 GCTGACCATGTTACTATGTTGGTCTCTTCTTCTGATGTCATGTCATGTCAGCAGCGTGTACA 399
Qy 858 TACAGGACATATGATGCTCGGGGAGGCGCCAGTGGTTCAGCAACGTTTGAACCGGTG 917
Db 400 TACCGCGCAACGATGCGCGAGGTGGCGCGGAATGGTCAACACGATTCGCTCTCTC 459
Qy 918 AACAGCTGGCGGATACGTTAACTGATATAAGCCGTCGATTTGCTGTGCCAGTCAAG 977
Db 460 GACAGCTGGCGCGGACCAACAGACCTGACAGGCCGCTGCTGCTGTGCCCATCAAG 519
Qy 978 AAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTCTCTGACTGGTAAATGTTGCC 1037
Db 520 CAAAAGTACGGCAGCAAGATCTCATGGCTGACCTTATCGTCTCTCGCGGCAACGTCGCC 579
Qy 1038 CTTGAATCCTATGGGATTTAAACCGCTGGATTTCTGTGGCGGAAGAAGATGACTGGGAG 1097
Db 580 CTCGAGCACAGCGGCTTCGAGACCCCTGGTTCGCGCGTGTGCGGCCGACACCTGGGAA 639
Qy 1098 TCAGA 1102
Db 640 GCCGA 644

RESULT 3

CD422961
LOCUS SAI_38_H10.gi_A002 Salicylic acid-treated seedlings Sorghum bicolor
DEFINITION cDNA clone SAI_38_H10_A002 5', mRNA sequence.
ACCESSION CD422961
VERSION CD422961.1 GI:31329224
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Sorghum.

1. (bases 1 to 701)
 Cordonnier-Pratt.M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
 Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Salzman,R.,
 Chua Tan,N., Gonzalez,M., Lane,S., Miller,V., Nanda,P.,
 Olaseinde,O., Eastman,A. and Pratt,L.H.
 An EST database from Sorghum: salicylic acid-treated seedlings
 Unpublished (2003)
 Other_ESTs: SAL_38_H10.b1_A002
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
 the Human Genome Center, University of Tokyo Institute of Medical
 Science; plant material and RNA prepared at Texas A & M University;
 sequencing done in the Laboratory for Genomics and Bioinformatics,
 University of Georgia. Sequence ends have been trimmed to exclude
 vector and regions below Phred quality 16. Three-prime sequences
 are presented as their reverse complement and have been trimmed to
 exclude polyA.

Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).

FEATURES

source
 1..701
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultiivar="BTx623"
 /db_xref="taxon:4558"
 /clone="SAL_38_H10_A002"
 /lab_host="DH10B-Ti phage-resistant E. coli"
 /clone_lib="Salicylic acid-treated seedlings"
 /note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
 library was prepared from polyA+ RNA from seedlings grown
 in hydroponic culture. At 8 days of age, medium was
 supplemented with 1 mM salicylic acid (SA). Roots and
 shoots were harvested after 27 and 72 hr and material from
 both time points was combined prior to RNA isolation.
 Double-stranded cDNA was cloned unidirectionally into
 different DraIII sites of the pME18S-FL3 vector (5'-prime
 DraIII site is CACTGTTGTG, 3'-prime DraIII site is
 CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 13.4%; Score 200; DB 6; Length 701;
 Best Local Similarity 66.8%; Pred. No. 8.6e-50;
 Matches 284; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
 Qy 678 CCTGAATCAATCCCTGGGGGCTGATTTTGAATATGCGACAGATTTCAACAGCTGGAT 737
 Db |||||
 Qy 264 CCTCAGTATACCCATTCGGTGTCTTCATTCGACTACACCGAGGCTTTCAAGTCTCTGGAC 323
 Db |||||
 Qy 738 ATGGAGGCTCTGAARAAGATATCAAGATTTCGTGACAACTTCCAGATTTGGTCCCT 797
 Db |||||
 Qy 324 TTGCTGATTTGAAGAAGACCTCAACGCGTTCTCACTGATTTCCAGATTTGGTGGCT 383
 Db |||||
 Qy 798 CGGATATGTCATTATGTCCTTCTTTATTTCGTATGGTTCGACGGTCCCGGAACA 857
 Db |||||
 Qy 384 GCTGACCATGTAATGATGTTGCTCTCTTCATTCGTATGTCATGCAAGCGCTGTGACA 443
 Db |||||
 Qy 858 TACAGGACATATGATGCGGGGAGCGCCAGTGGTGTGTCAGCAACGTTTTGAACCGGTG 917
 Db |||||
 Qy 444 TACCGCGCAATGATGCGCGAGGTGGCGCGGATGGGTCAACACGATTCGCTCTCTC 503
 Db |||||
 Qy 918 AACAGCTGGCGGATTAAGTTAATCTGATAAAGCCCTGATTCCTGTGCGCAGTCAAG 977
 Db |||||
 Qy 504 GACAGCTGGCGGACAAACAGAACTGGCAAGGCCGCTGCTCTGCTGCGCCCATCAAG 563
 Db |||||
 Qy 978 AAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTCTGACCTGGTAAATGTTGCC 1037
 Db |||||
 Qy 564 CAAAAGTACGCGACGAGATCTATGGGCTGACCTTTATCGTCTCGCGCGCAACGTCGCC 623
 Db |||||

Qy 1038 CTTGAATCCATGGGATTTAAACCGTGGGATTTCGTGCGGAGAGAGATGACTGGGAG 1097
 Db |||||
 Qy 624 CTCGAGCACAGCGGCTTCGAGACCCCTCGGTTTCGCGGTGTCGCGCGCACACTGNGAA 683
 Db |||||
 Qy 1098 TCAGA 1102
 Db 684 GCCGA 688
 Db |||||

RESULT 4

BO752115
 LOCUS BO752115
 DEFINITION ESF632678 DSCT Colletotrichum trifolii cDNA clone pDSCT10-12, mRNA
 ACCESSION BO752115
 VERSION BO752115.1 GI:21907520
 KEYWORDS EST.
 SOURCE Colletotrichum trifolii
 ORGANISM Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 REFERENCE 1 (bases 1 to 762)
 AUTHORS Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
 Cheung,F. and Fraser,C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other_ESTs: EST632677
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbya@puccini.crl.umn.edu
 TIGR sequence name: MTSAJ12TV More information is available at:
 www.medicago.org

Seq primer: (gtA ATA Cga Ctc Act ATA 999 C).

FEATURES

source
 1..762
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSCT10-12"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after
 inoculation) grown in liquid culture (cutin minimal medium
 containing 2%glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
 EcoRI; isolate: 2sp2; cDNA was prepared from polyA+
 enriched RNA The cDNA was ligated into lambda gcll from
 Stratagene and packaged using Gigapack packaging extracts.
 An aliquot of the amplified library was used to transduce
 E. coli Y1090 and phage DNA was purified from a liquid
 lysate. The cDNA inserts were gel purified after EcoRI
 digestion and ligated into pBluescript SK+. Aliquots of
 the ligation were used to transform E. coli DH5alpha which
 were plated onto medium with X-gal for selection of
 recombinants."

ORIGIN

Query Match 13.2%; Score 197.2; DB 5; Length 762;
 Best Local Similarity 62.2%; Pred. No. 6.5e-49;
 Matches 310; Conservative 0; Mismatches 188; Indels 0; Gaps 0;
 Qy 632 CTACTATCCAGAACACTGGATTTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
 Db |||||
 Qy 152 CTGTGCGCGGATCGCTGAAGCTCAACATCCTCCGCGACGACACGCCCTGCACCAACC 211
 Db |||||

COMMENT

Other ESTs: EST632074
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Boriang Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbyepuccini.crl.umn.edu
TIGR sequence name: MTSAE66TV More information is available at:
www.medicago.org
Seq primer: (gta Ata Cga Ctc Act Ata 999 C).

FEATURES

source
1..639
Location/Qualifiers
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSC75-66"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DHSalpha"
/clone_lib="DSC7"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 28p2; cDNA was prepared from polyA+ enriched RNA the cDNA was ligated into lambda gill from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DHSalpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 12.8%; Score 191; DB 5; Length 639;
Best Local Similarity 64.2%; Pred. No. 5e-47;
Matches 287; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
Qy 632 CTACTATCCAGAAACACTGGATTAACTCTCTGAGATTACAGACCCCTGAATCAATCC 691
Db 182 CTGTGGCCGATGGCTGAGCTCAACATCTTCGCCAGACACGCCCGCTACCAACCC 241
Qy 692 CTGGGGGCTGATTTGATATGCCACAGATTTCAACAGCTGGATATGGAGGCTCTGAA 751
Db 242 GCTGGCCAGGACTTTGACTACGTCCGCCCTTCAAGTCGCTCGACTACGAGGGCGTCAA 301
Qy 752 AAAAGATATCAAGATTGCTGACAACTTCCAGGATGGTGGCCCTGGGATTTATGTCA 811
Db 302 GAAGGACCTCAGGCCCTGTGATGACCGACTCCAGGACTGGTGGCTGCGGACTTTGGCCA 361
Qy 812 TTATGGTCTTTCTTTATTCGTATGGCTTGGCAGGTCGCCGGAACATACAGGACATATGA 871
Db 362 CTACGGGGCTGTGTTTCATCCCATGGCTGGCAGCGCCGCGACGTCACGATTCACGA 421
Qy 872 TGGCCGGGGAGCGCCAGTGGTGTGTCAGCAACGTTTGAACCGCTGAACAGCTGGCCGGA 931
Db 422 CGGACGCGGAGTGTGTGAGAGGGCCAGCAACGCTTCGCACCGCTCAACAGCTGGCCGGA 481
Qy 932 TAAGCTTAATCTGATTAAGCCCGTCAATTGCTGTGGCCAGTCAAGAAAAAATACGGCTC 991
Db 482 CAATGTGAGCCCTGACAAAGGCCCGCTCGGCTCTGTGGCCATCAAGAAAAAGTACGGCAA 541
Qy 992 CAGTATTTCTGGGGAGACCTGATGGTCCGCTGACCTGATTAATGTTGCCCTTGAATCCATGG 1051
Db 542 CAAGATCTGTGGGCCGACCTCATGATCCTGGCCGCAACGCTGGCCCTCGAGTCCATGGG 601
Qy 1052 ATTTAAACGCTGGGATTTGCTGGCGG 1078
Db 602 TTTCACAGCGGGCGCTTCTCCGAGG 628

RESULT 7

BZ893918/c
LOCUS
DEFINITION
HL8 0153 H1 pUC18 Library Halorubrum lacusprofundi genomic 5',
genomic survey sequence.
ACCESSION
BZ893918
VERSION
BZ893918.1 GI:33344508
KEYWORDS
GSS.
SOURCE
Halorubrum lacusprofundi
ORGANISM
Halorubrum lacusprofundi
Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
Halobacteriaceae; Halorubrum.
REFERENCE
1 (bases 1 to 600)
AUTHORS
Goo, Y., Roach, J., Glusman, G., Baliga, N.S., Deutsch, K., Fan, M.,
Dasarma, S., Ng, W.V. and Hood, L.
TITLE
Low-pass Sequencing for Microbial Comparative Genomics
JOURNAL
Unpublished (2003)
COMMENT
Contact: Goo Y
Institute for Systems Biology
1441 North 34th Street, Seattle, WA 98103, USA
Tel: 206 732 1412
Fax: 206 732 1299
Email: ygoosystemsbiology.org
Seq primer: M13 Forward
Class: Shotgun.
Location/Qualifiers
1..600
/organism="Halorubrum lacusprofundi"
/mol_type="genomic DNA"
/strain="ATCC 49239"
/db_xref="taxon:2247"
/clone_lib="H1 pUC18 Library"
/note="Vector: pUC18; Site 1: SmaI; A shotgun library was
constructed from Halorubrum lacusprofundi genomic DNA
using pUC18/SmaI/BAP plasmid"

ORIGIN

Query Match 12.8%; Score 190; DB 8; Length 600;
Best Local Similarity 66.2%; Pred. No. 1e-46;
Matches 274; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 690 CCCTGGGGGCTGATTTGATTTATGATTTGCCACAGATTTCAACAGCTGGATATGGAGGCTCTG 749
Db 530 CCGATGGACGAGGAGTTTCGACTACGCCGCGCGTTTCGAGGAGCTCGACCTCGATCGCGTG 471
Qy 750 AAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTCGCGATTTATGGT 809
Db 470 AAACCGGATATCGAGGACGTAATGACGACATCGCAGGACTGGTGGCCGCGACTACGGC 411
Qy 810 CATTATGCTCTTTCTTTTATTTGATGCTGCTGCGTGGCAGCATACAGGACATAT 869
Db 410 ACCTACGGCGCGCTTTTTCATCCGGATGGCGTGGCAGCGCGGTAGTACCGGACGCAC 351
Qy 870 GATGCGCGGGAGCGCCAGTGGTGTGTCAGCAACGTTTTGAAACCGCTGAAACAGCTGGCCG 929
Db 350 GACGCGACGCGCGCGCGCGCGCGCGGACGACGCGCTCCACCGCTTAAACAGCTGGCG 291
Qy 930 GATAACGTTAATCTGATTAAGCCCGCTCGATTGCTGTGGCCAGTCAAGAAAAAATACGGC 989
Db 290 GACACGCTCAACCTCGCAAGGCGCGACGACTGCTGTGGCCGCTCAACAGAGTACGGT 231
Qy 990 TCAGATATTTCTGGGGAGACCTGATGGTCTGCTGATGTAATGTTGCCCTTGAATCCATG 1049
Db 230 CGCAAGCTCTCTGTGGCGGATCTAATGTTCTACTGGGAACGTCGCGCTCGAGTCGATA 171
Qy 1050 GGATTTAAACAGCTGGGATTTGTCGCGGAAGAGAAGATCACTGGGAGTCCGAC 1103
Db 170 GGATTCGAGACGTTGGCTTCGGCGCGCGCGGAGGACGAGTTCAAGCCCGAC 117

RESULT 8

BQ751881

```

LOCUS       BQ751881               742 bp    mRNA    linear    EST 18-JUL-2002
DEFINITION   EST632444 DSCT Colletotrichum trifolii cDNA clone pDSCT8-27, mRNA
sequence.
ACCESSION   BQ751881
VERSION     BQ751881.1   GI:21907286
KEYWORDS    EST.
SOURCE      Colletotrichum trifolii
ORGANISM    Colletotrichum trifolii
REFERENCE   1 (bases 1 to 742)
AUTHORS     Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
            Cheung, F. and Fraser, C.M.
TITLE       ESRs from mycelia of Colletotrichum trifolii race 1
JOURNAL     Unpublished (2002)
COMMENT     Contact: Deborah A. Samac
            Department of Plant Pathology
            University of Minnesota
            495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
            Tel: 612 625 1243
            Fax: 651 649 5058
            Email: debbys@puccini.crl.umn.edu
            TIGR sequence name: MTSAH27TV More information is available at:
            www.medicago.org
            Seq primer: (gtA AtA CgA CtC Act AtA 999 C).
            Location/Qualifiers
                1..742
                    /organism="Colletotrichum trifolii"
                    /mol_type="mRNA"
                    /strain="race 1"
                    /db_xref="taxon:5466"
                    /clones="pDSCT8-27"
                    /tissue_type="mycelia"
                    /dev_stage="Young, actively growing mycelia (3 days after
                    inoculation) grown in liquid culture (cutin minimal medium
                    containing 2%glucose)."
                    /lab_host="DH5alpha"
                    /clone_lib="DSCt"
                    /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                    EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+
                    enriched RNA. The cDNA was ligated into Lambda gt11 from
                    Stratagene and packaged using Gigapack packaging extracts.
                    An aliquot of the amplified library was used to transduce
                    E. coli Y1090 and phage DNA was purified from a liquid
                    lysate. The cDNA inserts were gel purified after EcoRI
                    digestion and ligated into pBluescript SK+. Aliquots of
                    the ligation were used to transform E. coli DH5alpha which
                    were plated onto medium with X-gal for selection of
                    recombinants."

ORIGIN
Query Match      12.4%; Score 185.2; DB 5; Length 742;
Best Local Similarity 62.0%; Pred. No. 3.4e-45;
Matches 310; Conservative 0; Mismatches 188; Indels 2; Gaps 1;

Qy      632 CTACTATCCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGTAATCAATCC 691
Db      178 CTGTGGCCGATGCGCTGAAGCTCAACATCTCTCGGCAGCACACGCCGCTGACCAACC 237
Qy      692 CTGGGGGCGCTGATTTTGATTATGCCACAGATTTCACAGCTGGATATGGAGGCTCTGAA 751
Db      238 GCTGGCCAGGACTTTTGACTACTAGTCGCGCCCTTCAAGTCGCTACGAGGGCGTCAA 297
Qy      752 AAAAGATATCAAGATTGCTGACAACTTCCAGGATTGGTGGCCCTGGGATTATGGTCA 811
Db      298 GAAGACTCTACGCCCTGATGACCGACTCCAGGACTGGTGGCTGCGGACTTTGGCCA 357
Qy      812 TTATGGTCTTTCTTTATTCGATGGCTTGGCAGCGTGGCCGGAACATACAGACATATGA 871
Db      358 CTACGGCGGTCTGTTTCATCCGATGGCTGGCACAGCGCCGCGACGTTACCGAGTTACGA 417

```

```

Qy      872 TGCCCGGGGA--GGCCCCAGTGGTGTGTCAGCAACGTTTGAACCGCTGAACAGCTGGCCG 929
Db      418 CGGACCGCGAGGGGGGTGGGAGAGGGCCAGCAACGCTTCGACACCGCTCAACAGCTGGCCG 477
Qy      930 GATAACGTTAATCTGATAAAGCCCTCGATTGCTGTGGCCACTCAAGAAAAAATACGGC 989
Db      478 GACAATGTGAGCTTCGACCAAGGCCCTCGGCTCTGTGGCCCATCAAGCAAAAGTAGTCGGC 537
Qy      990 TCCAGTATTTCCTGGGGAGACCTGATGTCCTGACTGGTAATGTTCCTTGAATCCATG 1049
Db      538 AACAGATCTCTGTGGCCGACCTCATGATCTCTGGCCGCAACGTGGCCCTCGAGTCCATG 597
Qy      1050 GGATTTAAACGCTGGGATTGTCGCGGAAGAAAGATGACTGGGAGTGGAGCTCGTA 1109
Db      598 GGTTTCCAGAGCGCGGCTTCTCCGGAGGCGCTCCGACACCTGGGAAGCCGACGAGTCC 657
Qy      1110 TACTGGGGGCTGACAACAA 1129
Db      658 GTCTACTGGGGCGCGGAGAA 677

RESULT 9
BQ751387
LOCUS     BQ751387
DEFINITION EST631950 DSCT Colletotrichum trifolii cDNA clone pDSCT4-87, mRNA
sequence.
ACCESSION BQ751387
VERSION   BQ751387.1   GI:21906792
KEYWORDS  EST.
SOURCE    Colletotrichum trifolii
ORGANISM  Colletotrichum trifolii
REFERENCE 1 (bases 1 to 606)
AUTHORS   Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
            Cheung, F. and Fraser, C.M.
TITLE     ESRs from mycelia of Colletotrichum trifolii race 1
JOURNAL   Unpublished (2002)
COMMENT   Other_ESTRs: EST631951
            Contact: Deborah A. Samac
            Department of Plant Pathology
            University of Minnesota
            495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
            Tel: 612 625 1243
            Fax: 651 649 5058
            Email: debbys@puccini.crl.umn.edu
            TIGR sequence name: MTSAD87TK More information is available at:
            www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gTg gAT CC) .
            Location/Qualifiers
                1..606
                    /organism="Colletotrichum trifolii"
                    /mol_type="mRNA"
                    /strain="race 1"
                    /db_xref="taxon:5466"
                    /clones="pDSCT4-87"
                    /tissue_type="mycelia"
                    /dev_stage="Young, actively growing mycelia (3 days after
                    inoculation) grown in liquid culture (cutin minimal medium
                    containing 2%glucose)."
                    /lab_host="DH5alpha"
                    /clone_lib="DSCt"
                    /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                    EcoRI; isolate: 2sp2 ; cDNA was ligated into Lambda gt11 from
                    Stratagene and packaged using Gigapack packaging extracts.
                    An aliquot of the amplified library was used to transduce
                    E. coli Y1090 and phage DNA was purified from a liquid
                    lysate. The cDNA inserts were gel purified after EcoRI
                    digestion and ligated into pBluescript SK+. Aliquots of
                    the ligation were used to transform E. coli DH5alpha which
                    were plated onto medium with X-gal for selection of
                    recombinants."

```

ORIGIN	recombinants."	
Query Match	12.3%;	Score 183.2; DB 5; Length 606;
Best Local Similarity	63.1%;	Pred. No. 1.3e-44;
Matches	298; Conservative	0; Mismatches 173; Indels 1; Gaps 1;
Qy	632	CTACTATCCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGATCAATCC 691
Db	135	CTGTGGCGGATGGCTGAAGCTCAACATCTCGCCAGCACAGCCGTCACAAACC 194
Qy	692	CTGGGGGCGCTGATTTTGATTATGCCACAGATTTCACAGCTGGATATGGAGGCTCGAA 751
Db	195	GCTGGCCAGGACTTTGACTACGTGCGCGCTTCAAGTCGCTCGACTACGAGGGCGTCAA 254
Qy	752	AAAAGATATCAAGATTGCTGACAACTTCCAGGATTTGGTCCCTGGGATATATGGTCA 811
Db	255	GAAGGACCTCACGGCCCTGATGACCGATCTCCAGGACTGGTGGCTGCGGACTTTGGCCA 314
Qy	812	TTATGGTCTCTTTCTTTATTCGTATGGCTTGGCAGGTCGCCGGAACATACAGCATATGA 871
Db	315	CTACGGGGCTGTGTATCCCATGGCTTGGCAGCGCCGACAGTACCGAGTTCACGA 374
Qy	872	TGGCCGGGGAGG-CGCCAGTGGTGTGTCAGCAACGTTTGAACCGCTGAACAGCTGGCGG 930
Db	375	CGGACGGGGAGTGGTGGAGAGGGCCNAACGCTTCGCACCGCTCAACAGCTGGCGG 434
Qy	931	ATAAGTTATCTGGATAAAGCCGCTGATTTGCTGTGGCCAGTCAAGAAAAAATACGGCT 990
Db	435	ACAATGTACGCTCGACAAAGCCGCTGGCTGTGTGGCCCATCAAGCAAAAGTACGGCA 494
Qy	991	CCAGTATTTCTGGGGAGACCTGATGTCCTGACTGATGTTGCTTATGTTGCCCTTGAATCCATGG 1050
Db	495	ACAAGATCTCTGGGGCGGACCTCATGATCTCTGGCGGCAAGCTGGCCCTCGAGTCCATGG 554
Qy	1051	GATTAAACAGCTCGGATTTGCTGGCGGAAGAGATGACTGGGAGTCGGA 1102
Db	555	GTITCCAGACGGCGGCTTCTCCGAGGCGCTCCGACACCTTGGGAAGCCGA 606
RESULT 10		
CF876888		
LOCUS	CF876888 726 bp mRNA linear EST 31-OCT-2003	
DEFINITION	trio074xa08.b1 T.reesei mycelial culture, Version 6 October 2003	
ACCESSION	Hypocrea jecorina cDNA clone trio074xa08, mRNA sequence.	
VERSION	CF876888.1 GI:38131570	
KEYWORDS	EST.	
SOURCE	Hypocrea jecorina (anamorph: Trichoderma reesei)	
ORGANISM	Hypocrea jecorina	
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.	
AUTHORS	Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D., Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K., Ward, M. and Dean, R.A.	
TITLE	Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from Trichoderma reesei	
JOURNAL	FEMS Microbiol. Lett. 230 (2), 275-282 (2004)	
COMMENT	Contact: Ralph A. Dean Fungal Genomics Laboratory North Carolina State University Campus Box 7251, Raleigh, NC 27695, USA Tel: 919-513-0020 Fax: 919-513-0024 Email: ralph.dean@ncsu.edu Seq primer: LT-F1 primer. Location/Qualifiers 1. .726 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="QM6a"	
FEATURES		
Source		

ORIGIN		
Query Match	12.3%;	Score 183; DB 7; Length 726;
Best Local Similarity	60.1%;	Pred. No. 1.6e-44;
Matches	303; Conservative	0; Mismatches 201; Indels 0; Gaps 0;
Qy	625	AAATTTCTACTATCCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAT 684
Db	219	ACRAGGACTGGTGGCCGAGTCCCTCAAGCTCAACATCTCCGCCAGCACACAGCGTCA 278
Qy	685	CAATCCCTGGGGGCTGATTTTGATTATGCCACAGATTTCACAGCTGGATATGGAGG 744
Db	279	CAATCTCTCGGCCCGGACTTTTGACTACCGGGTGGCTTCAAGACCTTCGACTACGATG 338
Qy	745	CTCTGAAAAAGATATCAAGATTTCGTGACAACTTCCAGGATTTGGTCCCTCGGAT 804
Db	339	CCGTCAAGAAAGGACTTTACGGCCCTGATGACCGATTCTCAGGACTGGTGGCCCGGACT 398
Qy	805	ATGTCATATATGTCCTTTCTTTATTCGTATGGCTTGGACGGTGGCCGGAACATACAGGA 864
Db	399	TTGGCCACTACGGAGGCTCTTTATCCGGCTGGCATGGCACAGCGTGGCACGATATCGTG 458
Qy	865	CATATGATGCCCCGGGAGCGCCAGTGGTGGTCAAGCAACGTTTGAACCCGTGAACAGCT 924
Db	459	TCACTGACGGCCGCTGGCGGCGGTGGCCAGGCGCCAAACGATTCGCACCTCTCAACAGCT 518
Qy	925	GGCCGGAATAACGTTTAACTGATTAAGCCGCTGATTGCTGTGGCCAGTCAAGAAAAAT 984
Db	519	GGCCGGAATAACGTCAGCTTGGACAGGCTCGCCGTCTGCTGTGGCCCATCAAGCAAAAGT 578
Qy	985	ACGCTCCAGATATTTCTGGGGAGACCTGATGTCCTGACTGTGTAATGTTGCCCTTGAAT 1044
Db	579	ACGGCAACAAGCTCTCTGGGGCGGACCTCTCATCTCTCCGCAACGCTGCTCGAGT 638
Qy	1045	CCATGGATTTAAACCGCTGGGATTTGCTGGCGGAAGAGATGACTGGGAGTCGGACC 1104
Db	639	CCATGGCTTCAAGACCGCTCGGATTTCCGGCGGCGNCCTCCAGACACTTGGAGGCTGACG 698
Qy	1105	TGGTATATCTGGGGCGCTGCAACA 1128
Db	699	AGGCTGCTACTGGGGCGGCGAGA 722
RESULT 11		
CB905354		
LOCUS	CB905354 780 bp mRNA linear EST 02-JUL-2003	
DEFINITION	trio074xa08 T.reesei mycelial culture, Version 3 april Hypocrea	
ACCESSION	jecorina cDNA clone trio074xa08, mRNA sequence.	
VERSION	CB905354	
KEYWORDS	EST.	
SOURCE	Hypocrea jecorina (anamorph: Trichoderma reesei)	
ORGANISM	Hypocrea jecorina	
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.	
AUTHORS	Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J., Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C., Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M.	
TITLE	Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei	
JOURNAL	J. Biol. Chem. 278 (34), 31988-31997 (2003)	
MEDLINE	22803314	
PUBMED	12788920	

RESULT 13
CF870162 726 bp mRNA linear EST 31-OCT-2003
LOCUS trico21x120.b1 T.reesei mycelial culture, Version 6 October 2003
DEFINITION Hypocrea jecorina cDNA clone trico21x120, mRNA sequence.
ACCESSION CF870162
VERSION CF870162.1 GI:38124844
KEYWORDS EST.
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE 1 (bases 1 to 726)
AUTHORS Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D., Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K., Ward, M. and Dean, R.A.
TITLE Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from Trichoderma reesei
JOURNAL FEMS Microbiol. Lett. 230 (2), 275-282 (2004)
COMMENT Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu
Seq primer: LT-F1 primer.
Location/Qualifiers
1. 726
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clones="trico21x120"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pREP3Y; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
ORIGIN
Query Match 12.2%; Score 182.4; DB 7; Length 726;
Best Local Similarity 60.1%; Pred. No. 2.4e-44;
Matches 303; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
Qy 625 AAAATTTCTACTATCCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAAT 684
Db 211 ACAAGGACTGTGCGCCGAGTCCCTCAAGCTCAACATCTCTCGGCAGACACACAGTCA 270
Qy 685 CAAATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATATGGAGG 744
Db 271 CCAATCTCTCGGCCCGGACTTTGACTACCCGGCTGCCCTTCAAGACCCTCGACTACGATG 330
Qy 745 CTCTGAAAAGATATCAAGATTTGCTGACAACTCCACAGATTTGGCCCTCGGGATT 804
Db 331 CCGTCAAGAGAGGACCTTACCGCCCTGATGACCGATTCTCAGGACTGGTGGGCCCGGACT 390
Qy 805 ATGCTGATTATGGTCTTTCTTTTATTCGTATGGCTTGGCAGCGTGCCGGAACATACAGGA 864
Db 391 TTGGCCACTACGAGGCTCTTTATTCGGCTGGGATGACAGCGCTGGACGCTATCGTG 450
Qy 865 CATATGATGGCCGGGAGGCGCAGTGGTGTGACCAACGTTTTTGAACCGCTGAACAGCT 924
Db 451 TCACTGACGGCCGTGGGGCGGTGGCCAGGCGCAACACGATTTCGACCTCTCAACAGCT 510
Qy 925 GGCCGGATAAGTTATCTGATTAAGCCCTCGATTGCTGTGGCAGTCAAGAAAAAT 984
Db 511 GGCCCGCAACCTGAGCTTGACAAAGCTCGCCGCTGCTGTGGCCCATCAAGCAAGT 570
Qy 985 ACGGCTCAGTATTTCTGGGGAGACCTGATGGTCTGACCTGGTAATCTTCCCTTGAAT 1044

Db 571 ACGCAACAAGCTCTCGTGGGCGGACCTCTCATCTCTCCGGCAACGTTGGCTCTCGAGT 630
Qy 1045 CCATGGGATTTAAACCGCTGGGATTTGCTGGGGAAGAAGATGACTGGAGTGGAGCC 1104
Db 631 CCATGGGCTTCAAGACGCTCGGATTTGCGCGCGCGCTCCAGACACTTTGGGAGGCTGACG 690
Qy 1105 TGGTATATCTGGGGCCCTGACAACA 1128
Db 691 AGGCTGCTACTGGGGCGGAGG 714
RESULT 14
CF869494 748 bp mRNA linear EST 31-OCT-2003
LOCUS trico19xi07.b1 T.reesei mycelial culture, Version 6 October 2003
DEFINITION Hypocrea jecorina cDNA clone trico19xi07, mRNA sequence.
ACCESSION CF869494
VERSION CF869494.1 GI:38124176
KEYWORDS EST.
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE 1 (bases 1 to 748)
AUTHORS Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D., Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K., Ward, M. and Dean, R.A.
TITLE Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from Trichoderma reesei
JOURNAL FEMS Microbiol. Lett. 230 (2), 275-282 (2004)
COMMENT Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu
Seq primer: LT-F1 primer.
Location/Qualifiers
1. 748
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clones="trico19xi07"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October 2003"
/note="Vector: pREP3Y; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."
ORIGIN
Query Match 12.2%; Score 182.4; DB 7; Length 748;
Best Local Similarity 60.1%; Pred. No. 2.5e-44;
Matches 303; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
Qy 625 AAAATTTCTACTATCCAGAAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAAT 684
Db 213 ACAAGGACTGTGCGCCGAGTCCCTCAAGCTCAACATCTCTCGGCAGACACACAGTCA 272
Qy 685 CAAATCCCTGGGGGCTGATTTGATTATGCCACAGATTTCAACAGCTGGATATGGAGG 744
Db 273 CCAATCTCTCGGCCCGGACTTTGACTACCCGGCTGCCCTTCAAGACCCTCGACTACGATG 332
Qy 745 CTCTGAAAAGATATCAAGATTTGCTGACAACTCCACAGATTTGGCCCTCGGGATT 804
Db 333 CCGTCAAGAGAGGACCTTACGGCCCTGATGACCGATTCTCAGGACTGGTGGGCCCGGACT 392
Qy 805 ATGCTGATTATGGTCTTTCTTTTATTCGTATGGCTTGGCAGCGTGCCGGAACATACAGGA 864
Db 393 TTGGCCACTACGAGGCTCTTTATTCGGCTGGGATGACAGCGCTGGCAGCTATCGTG 452

```
Qy 865 CATATGATGCGCGGAGGCGCAGTGGTGTGCTCAGCAACGTTTGAACCGCTGAACAGCT 924
Db 453 TCACGTAGCGCGGTGGCGCGGTGGCCAGGCCAACCAACGATTCGACACTCTCAACAGCT 512
Qy 925 GGCCGGATAACGTTAAATCTGGATAAAGCCCGTCGATTTGCTGTGGCCAGTCAAGAAAAAAT 984
Db 513 GGCCCGACAACGTGAGCTTGGACAAGGCTGCGCGTCTGCTGTGGCCCATCAAGCAAAAGT 572
Qy 985 ACGGCTCAGTATTTCTGGGGAGACCTGATGGTTCCTGACCTGGTAATGTGGCCCTTGAAT 1044
Db 573 ACGGCAACAAGCTCTCGTGGGCCGACCTCTCATCTCTCGGCAACGCTGCTCTCGAGT 632
Qy 1045 CCATGGGATTTAAACCGTGGGATTTGTCGCGGAAGAGATGACTGGGAGTCGGACC 1104
Db 633 CCATGGGCTTCAGACGCTCGGATTTGCGCGGGCGTCCAGACACTTGGAGGCTGACG 692
Qy 1105 TGGTATATCTGGGGGCTGACAAACA 1128
Db 693 AGGCTGCTACTGGGCGCGGAGA 716

RESULT 15
CB899177
LOCUS
DEFINITION
    trico17xj21 T. reesei mycelial culture, Version 3 april Hypocrea
    jecorina cDNA clone trico17xj21, mRNA sequence.
ACCESSION
CB899177
VERSION
CB899177.1 GI:30113835
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
1 (bases 1 to 776)
Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J.,
Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C.,
Olivares,H.A., Teunissen,P.J., Yao,J. and Ward,M.
Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
JOURNAL
MEDLINE
PUBMED
22803314
COMMENT
Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-P1 primer.
Location/Qualifiers
    /organism="Hypocrea jecorina"
    /mol_type="mRNA"
    /strain="QM6a"
    /db_xref="taxon:51453"
    /dev_stages="mycelia"
    /clone_lib="T.reesei mycelial culture, Version 3 april"
    /note="Vector: pREP3y; Site 1: Not I/Sal I; Mycelial
    culture grown from 24 hrs to 6 days with varying Carbon
    and Nitrogen sources and concentrations."
FEATURES
    source
    1..776
ORIGIN
Query Match 12.2%; Score 182.4; DB 6; Length 776;
Best Local Similarity 60.1%; Pred. No. 2.5e-44;
Matches 303; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
Qy 625 AAAATTCTACTATCCAGAAACACTGGGATTTAACTCTCTGAGATTACAGAGCCCTGAT 684
Db 269 ACAAGGACTGGTGGCCCGAGTCCCTCAAGCTCAACATCTCCGCCGACACACAGACGTCA 328
```

```
Qy 685 CAATCCCTGGGGGCTGATTTTGTATATGCCACAGATTTCAACAGCTGGATATGGAG 744
Db 329 CCATCTCTCGGCCCGACTTTGACTACCCGGCTGCTTCAAGACCCCTCGACTACGATG 388
Qy 745 CTCTGAAAAAGATATCAAAAGATTTGCTGACACACTTCCCAAGGATTTGTCCTCGGATTT 804
Db 389 CCCTCAAGAAAGAGACCTTTACGGCCCTGATGACCCGATTTCTCAGGACTGGTGGCCCGGACT 448
Qy 805 ATGCTCATTTATGTCCTTTCTTTATTCGTATGGCTTGGCAGCGTGGCCGGAACATACAGGA 864
Db 449 TTGGCCACTACGAGGCGCTCTTTATCCGGCTGGCATGGCAGCGCTGGCACGTATCGTG 508
Qy 865 CATATGATGGCCGGGAGGCGCCAGTGGTGGTTCAGCAACGTTTTTGAACCGCTGAACAGCT 924
Db 509 TCACTGACGCGCGTGGCGCGGTGGCCAGGCCCAACAACGATTCGCACCTCTCAACAGCT 568
Qy 925 GGCCGGATAACGTTAAATCTGGATAAAGCCCGTCGATTTGCTGTGGCCAGTCAAGAAAAAAT 984
Db 569 GGCCCGACAACGTCAGCTTGGACAAGGCTGCGCGTCTGCTGTGGCCCATCAAGCAAAAGT 628
Qy 985 ACGGCTCAGTATTTCTGGGGAGACCTGATGGTTCCTGACTGGTAAATGTTGCCCTTGAAT 1044
Db 629 ACGCAACAAGCTCTCGTGGCCGACCTCTCATCTCTCCGGCAACGCTGCTCTCGAGT 688
Qy 1045 CCATGGGATTTAAACCGTGGGATTTTCTGTCGCGGAAGAGATGACTGGGAGTCGGACC 1104
Db 689 CCATGGGCTTCAGACGCTCGGATTTGCGCGGCCGCTCCAGACACTTGGGAGGCTGACG 748
Qy 1105 TGGTATATCTGGGGGCTGACAAACA 1128
Db 749 AGGCTGCTACTGGGCGCGGAGA 772

RESULT 16
CB900330
LOCUS
DEFINITION
    trico21x120 T. reesei mycelial culture, Version 3 april Hypocrea
    jecorina cDNA clone trico21x120, mRNA sequence.
ACCESSION
CB900330
VERSION
CB900330.1 GI:30114988
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
1 (bases 1 to 790)
Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J.,
Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C.,
Olivares,H.A., Teunissen,P.J., Yao,J. and Ward,M.
Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
JOURNAL
MEDLINE
PUBMED
22803314
COMMENT
Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-P1 primer.
Location/Qualifiers
    /organism="Hypocrea jecorina"
    /mol_type="mRNA"
    /strain="QM6a"
    /db_xref="taxon:51453"
    /clone_lib="T.reesei mycelial culture, Version 3 april"
    /note="Vector: pREP3y; Site 1: Not I/Sal I; Mycelial
    culture grown from 24 hrs to 6 days with varying Carbon
    and Nitrogen sources and concentrations."
FEATURES
    source
    1..790
```

and Nitrogen sources and concentrations."

ORIGIN

Query Match 12.2%; Score 182.4; DB 6; Length 790;
 Best Local Similarity 60.1%; Pred. No. 2.5e-44;
 Matches 303; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 625 AAAATTTCTACTATCCAGAACACATGGATTAACTCTCTGAGATTACACAGCCCTGAAT 684
 Db 275 ACAAGGACTGGTGGCCGAGTCCCTCAAGCTCAACATCTCCGCGACACACAGACGTC 334
 QY 685 CAAATCCCTGGGGGCTGATTGTGATATGACACAGATTTCAACAGCTGGATATGGAGG 744
 Db 335 CCAATCTCTCGGCCCGACTTTGACTACCGGCTGCCTTCAAGACCTCTGACTACGATG 394
 QY 745 CTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGAGATGGTGGCCCTGCGGAT 804
 Db 395 CCGTCAAGAGAGACCTTACGGCCCTGATGACCGATTTCTCAGGACTGGTGGCCCGGACT 454
 QY 805 ATGTCATATTATGTCCTTTCTTTATTTGATGATGACACAGATTTCAACAGCTGGATATGGAGG 864
 Db 455 TTGGCCACTACGAGGCTCTTTATCTCGGCTGGATGACACAGGCTGGCAGCTATCGTG 514
 QY 865 CATATGATGGCGGGGAGCGCCAGTGGTGTGCTGACCAACGTTTGAACCGCTGAACAGCT 924
 Db 515 TCATGACGGCGGTGGCGCGGTGGCCAGGCGCCACACAGATTCGACCTCTCACAGCT 574
 QY 925 GGCCGGATAACGTTAATCTGATAAAGCCCGTGCAGATTCGTGGCCAGTCAAGAAAAAT 984
 Db 575 GGCCCGCAACGCTGAGCTTGGACAAGGCTCGCCGTCTGCTGCGCCCATCAAGCAAAAGT 634
 QY 985 ACGCTCCAGTATTTCTGGGAGACCTGATGGTCTGCTGACGTTGATGTTGCCCTTGAAT 1044
 Db 635 ACGCAACAACGCTCTGCGGCGGACCTCTCATCTCTCGCGCAACGTTGCTCTCGAGT 694
 QY 1045 CCATGGGATTTAAACCGTGGGATTTGCTGCGGAAGAGAGATGACTGGGAGTCGGACC 1104
 Db 695 CCATGGGCTTCAAGACGCTCGGATTTGCGGCGGCGGCTCCAGACACTTGGGAGGCTGACG 754
 QY 1105 TGGTATACCTGGGCGCTGACAA 1128
 Db 755 AGGCTGCTACTGGGCGGCGAGA 778

RESULT 17

CB899636
 LOCUS
 DEFINITION
 jecorina cDNA clone tric019x107, mRNA sequence.
 EST.

ACCESSION
 CB899636
 VERSION
 CB899636.1 GI:30114294

KEYWORDS
 EST.

SOURCE
 Hypocrea jecorina (anamorph: Trichoderma reesei)

ORGANISM
 Hypocrea jecorina

REFERENCE
 1 (bases 1 to 803)
 Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,
 Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,
 Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,
 Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M.

TITLE
 Transcriptional regulation of biomass-degrading enzymes in the
 filamentous fungus Trichoderma reesei

JOURNAL
 J. Biol. Chem. 278 (34), 31988-31997 (2003)

MEDLINE
 22803314

PUBMED
 12788920

COMMENT
 Genencor Intl.

Contact: Pamela K. Foreman

925 Page Mill Road, Palo Alto, CA 94304, USA

Tel: (650) 846-7635

Fax: (650) 621-7817

Email: Pforeman@genencor.com

Seq primer: LT-F1 primer.

FEATURES
 source

Location/Qualifiers
 1..803
 /organism="Hypocrea jecorina"
 /mol_type="mRNA"
 /strain="QM6a"
 /db_xref="taxon:51453"
 /clone="tric019x107"
 /dev_stage="mycelia"
 /clone_lib="T.reesei mycelial culture, Version 3 april"
 /note="vector: pREP3; Site 1: Not I/Sal I; Mycelial
 culture grown from 24 hrs to 6 days with varying Carbon
 and Nitrogen sources and concentrations."

ORIGIN

Query Match 12.2%; Score 182.4; DB 6; Length 803;
 Best Local Similarity 60.1%; Pred. No. 2.6e-44;
 Matches 303; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 625 AAAATTTCTACTATCCAGAACACATGGATTAACTCTCTGAGATTACACAGCCCTGAAT 684
 Db 268 ACAAGGACTGGTGGCCGAGTCCCTCAAGCTCAACATCTCCGCGACACACAGACGTC 327
 QY 685 CAAATCCCTGGGGGCTGATTGTGATATGACACAGATTTCAACAGCTGGATATGGAGG 744
 Db 328 CCAATCTCTCGGCCCGACTTTGACTACCGGCTGCCTTCAAGACCTCTGACTACGATG 387
 QY 745 CTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGAGATGGTGGCCCTGCGGAT 804
 Db 388 CCGTCAAGAGAGACCTTACGGCCCTGATGACCGATTTCTCAGGACTGGTGGCCCGGACT 447
 QY 805 ATGTCATATTATGTCCTTTCTTTATTTGATGATGACACAGATTTGACAGCTGGCAGCTATCGTG 864
 Db 448 TTGGCCACTACGAGGCTCTTTATCCGGCTGGCATGGCACAGCGCTGGCACGCTATCGTG 507
 QY 865 CATATGATGGCGGGGAGCGCCAGTGGTGTGCTGACCAACGTTTGAACCGCTGAACAGCT 924
 Db 508 TCATGACGGCGGTGGCGCGGTGGCCAGGCGCCACACAGATTCGACCTCTCACAGCT 567
 QY 925 GGCCGGATAACGTTAATCTGATAAAGCCCGTGCAGATTCGTGGCCAGTCAAGAAAAAT 984
 Db 568 GGCCCGCAACGCTGAGCTTGGACAAGGCTCGCCGTCTGCTGCGCCCATCAAGCAAAAGT 627
 QY 985 ACGCTCCAGTATTTCTGGGAGACCTGATGGTCTGCTGACGTTGATGTTGCCCTTGAAT 1044
 Db 628 ACGCAACAACGCTCTGCGGCGGACCTCTCATCTCTCGCGCAACGTTGCTCTCGAGT 687
 QY 1045 CCATGGGATTTAAACCGTGGGATTTGCTGCGGAAGAGAGATGACTGGGAGTCGGACC 1104
 Db 688 CCATGGGCTTCAAGACGCTCGGATTTGCGGCGGCGGCTCCAGACACTTGGGAGGCTGACG 747
 QY 1105 TGGTATACCTGGGCGCTGACAA 1128
 Db 748 AGGCTGCTACTGGGCGGCGAGA 771

RESULT 18

CB901454

LOCUS

DEFINITION

tric026xk14 T.reesei mycelial culture, Version 3 april Hypocrea

jecorina cDNA clone tric026xk14, mRNA sequence.

ACCESSION

CB901454

VERSION

CB901454.1 GI:30116112

KEYWORDS

EST.

SOURCE

Hypocrea jecorina (anamorph: Trichoderma reesei)

ORGANISM

Hypocrea jecorina

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

REFERENCE

1 (bases 1 to 827)

AUTHORS

Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,
 Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,
 Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,
 Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M.

TITLE

Transcriptional regulation of biomass-degrading enzymes in the

JOURNAL
MEDLINE
PUBLISHED
COMMENT

filamentous fungus *Trichoderma reesei*
J. Biol. Chem. 278 (34), 31988-31997 (2003)
22803314
12788920

Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-F1 primer.

FEATURES

source
1. .827
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric026xk14"
/dev_stage="mycelia"
/clone_lib="T. reesei mycelial culture, Version 3 april"
/note="Vector: pREP3Y; Site_1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

ORIGIN

Query Match 12.2%; Score 181; DB 6; Length 827;
Best Local Similarity 60.3%; Pred. No. 7e-44;
Matches 298; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 625 AAAATTTCTACTATCCAGAAACACTGGATTTAATCTCTCTGAGATTACACAGCCCTGAAT 684
Db |||||
Qy 302 ACAGGACTGGTGGCCGAGTCCCTCAAGCTCAACATCTCCGCCAGCACACAGACGTCA 361
Db |||||
Qy 685 CAATATCCCTGGGGGGCTGATTTTGAATATGACCACAGATTTCAACAGCTGATATGAGG 744
Db |||||
Qy 362 CCAATCTCTCGGCCCGGACTTTGACTACCGGGTGCCTTCAAGACCCCTCGACTACGATG 421
Db |||||
Qy 745 CTCTGMAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTGGCCCTCGGATT 804
Db |||||
Qy 422 CCGTCAAGAAGACCTTACGGCCCTGATGACCGATTTCTCAGACTTGGTGGCCCGGACT 481
Db |||||
Qy 805 ATGTCATATTATGTCCTTTTCTTTATTCGTATGGCTTGGCAGCGTGGCCGAAACATACAGGA 864
Db |||||
Qy 482 TTGGCCACTACGAGGCTCTTTATCCGGCTGGCATGSCACAGCGTGGCACGTATCGTG 541
Db |||||
Qy 865 CATATATGCGGGAGGCGCCAGTGGTGTGACCAACGTTTGAACCGCTGAACAGCT 924
Db |||||
Qy 542 TCACTGACGGCGTGGCGCGGTGGCCAGGGCCCAACACGATTCGCACCTCTCAACAGCT 601
Db |||||
Qy 925 GGCCGGATAAGCTTAATCTGATTAAGCCCGTTCGATTGCTGGCCAGTCAAGAAAAAT 984
Db |||||
Qy 602 GGCCCGAACAGTGAAGTTGACAGAGGCTCCCGTCTCTGTGGCCCATCAAGCAAAAGT 661
Db |||||
Qy 985 ACGGCTCCAGTATTTCCTGGGGAGACCTGTATGGTCTGACTGGTAATTTGCCCTTGAAT 1044
Db |||||
Qy 662 ACGCAACAAGCTCTCGTGGCCGACCTCTCATCTCTCCGGCAACGTGCTCTCGAGT 721
Db |||||
Qy 1045 CCATGGGATTTAAACCGTGGGATTTGCTGGGGAAGAGATGATCGGAGTGGGACC 1104
Db |||||
Qy 722 CCATGGGCTTAAAGACGCTCGGATTTGCGGGCGCGTCCAGACACTTTGGGAGGCTGACG 781
Db |||||
Qy 1105 TGGTATACCTGGGG 1118
Db AGGTGNTACTGG 795

RESULT 19

CF871230
LOCUS
DEFINITION
ACCESSION
VERSION

CF871230.1
827 bp mRNA linear EST 31-OCT-2003
T. reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone tric026xk14, mRNA sequence.
CF871230
CF871230.1 GI:38125912

KEYWORDS

SOURCE
ORGANISM

Hypocrea jecorina (anamorph: *Trichoderma reesei*)
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

REFERENCE
AUTHORS

1. (bases 1 to 827)
Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,
Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K.,
Ward, M. and Dean, R.A.

TITLE

Characterization of the protein processing and secretion pathways
in a comprehensive set of expressed sequence tags from *Trichoderma*
reesei

JOURNAL

COMMENT

FEMS Microbiol. Lett. 230 (2), 275-282 (2004)

Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu
Seq primer: LT-F1 primer.

FEATURES

source

1. .827
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric026xk14"
/dev_stage="mycelia"
/clone_lib="T. reesei mycelial culture, Version 6 October
2003"

/note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

ORIGIN

Query Match 12.2%; Score 181; DB 7; Length 827;
Best Local Similarity 60.3%; Pred. No. 7e-44;
Matches 298; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

Qy 625 AAAATTTCTACTATCCAGAAACACTGGATTTAATCTCTCTGAGATTACACAGCCCTGAAT 684
Db |||||
Qy 302 ACAGGACTGGTGGCCGAGTCCCTCAAGCTCAACATCTCCGCCAGCACACAGACGTCA 361
Db |||||
Qy 685 CAATATCCCTGGGGGGCTGATTTTGAATATGACCACAGATTTCAACAGCTGATATGAGG 744
Db |||||
Qy 362 CCAATCTCTCGGCCCGGACTTTGACTACCGGGTGCCTTCAAGACCCCTCGACTACGATG 421
Db |||||
Qy 745 CTCTGMAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTGGCCCTCGGATT 804
Db |||||
Qy 422 CCGTCAAGAAGACCTTACGGCCCTGATGACCGATTTCTCAGACTTGGTGGCCCGGACT 481
Db |||||
Qy 805 ATGTCATATTATGTCCTTTTCTTTATTCGTATGGCTTGGCAGCGTGGCCGAAACATACAGGA 864
Db |||||
Qy 482 TTGGCCACTACGAGGCTCTTTATCCGGCTGGCATGSCACAGCGTGGCACGTATCGTG 541
Db |||||
Qy 865 CATATATGCGGGAGGCGCCAGTGGTGTGACCAACGTTTGAACCGCTGAACAGCT 924
Db |||||
Qy 542 TCACTGACGGCGTGGCGCGGTGGCCAGGGCCCAACACGATTCGCACCTCTCAACAGCT 601
Db |||||
Qy 925 GGCCGGATAAGCTTAATCTGATTAAGCCCGTTCGATTGCTGGCCAGTCAAGAAAAAT 984
Db |||||
Qy 602 GGCCCGAACAGTGAAGTTGACAGAGGCTCCCGTCTCTGTGGCCCATCAAGCAAAAGT 661
Db |||||
Qy 985 ACGGCTCCAGTATTTCCTGGGGAGACCTGTATGGTCTGACTGGTAATTTGCCCTTGAAT 1044
Db |||||
Qy 662 ACGCAACAAGCTCTCGTGGCCGACCTCTCATCTCTCCGGCAACGTGCTCTCGAGT 721
Db |||||
Qy 1045 CCATGGGATTTAAACCGTGGGATTTGCTGGGGAAGAGATGATCGGAGTGGGACC 1104
Db |||||
Qy 722 CCATGGGCTTAAAGACGCTCGGATTTGCGGGCGCGTCCAGACACTTTGGGAGGCTGACG 781
Db |||||
Qy 1105 TGGTATACCTGGGG 1118

[illegible]

/strains="ATCC 43049"
/db_xref="taxon:272569"
/clone_lib="Hm pUC18 Library"
/note="vector: pUC18; Site 1: SmaI; A shotgun library was
constructed from Haloarcula marismortui genomic DNA using
pUC18/SmaI/BAP plasmid"

ORIGIN

Query Match 11.9%; Score 177.4; DB 8; Length 720;
Best Local Similarity 61.5%; Pred. No. 8.6e-43;
Matches 324; Conservative 0; Mismatches 191; Indels 12; Gaps 2;
Qy 973 TCAAGAAAAATACGGCTCCAGTATTTCTCGGAGACCTGATGCTCTGCTGCTGTAATG 1032
Db 720 TCAAAAAAGAGTCCGCGCGCAAGCTCTGCTGGCGGACCTGATGCTGCTGCTGCAACC 661
Qy 1033 TTGCGCTTGAATCATGGGATTTAAACCTGGGATTTGCTGGCGGAGAGAAAGATGACT 1092
Db 660 ACGCATCGAATGATGGGGCTCAAGACGTTTCGGCTGGCGCGCGCGCGGAGATGCGT 601
Qy 1093 GGGAGTCGGAC---CTGGTATCTGGGGCTGACACAG-----CCTTTGCGAG 1140
Db 600 TTGAACCGGACGAGAGTGCAGTGGGGGCGCGGAGAGATGGAAGCCCATCAGTCCG 541
Qy 1141 ATAAACCGGATATAAAGCGGAAACTTCAGAAACCTCTTGGCGCCACGAGATGGGACTTA 1200
Db 540 AGCGCGTACTGACGAGCGGCACTCAAGAGCTCTGGGTGCTGCGGTGATGGGGCTGA 481
Qy 1201 TTTATGTCAATCTCGAAGGCCCGGTGGAACACAGATCTCTGGCTTCCGCGGAAAGATA 1260
Db 480 TTTACGTGGACCCAGAGGGGCGCAACCGCAACCGGACCGCTCGCGTCCGCGGAAATA 421
Qy 1261 TCAGGGAAGCTTTTACGATATGCGCATGCGATGATGAGGAGACTGTGCGCCCTGATCGGG 1320
Db 420 TCCGCGAGTCTGCGTCCGATGCGGATGAACGACGAGGAAACGCGGCACTCATCGCTG 361
Qy 1321 GAGGCGATACATTTGGTAAAGCAGATGTCGAGCGTCTCTGAAATATGATTGGCGCAG 1380
Db 360 CGGCGCACAGTTCGGGAGGTTCACGTCGCGGACGACCCGAGGAGAACCTCGGCGACG 301
Qy 1381 GGCTGTATGTGACCTGTGGAGGAGCAGGACTGGGATGGAAATAAATATGTTGTACAG 1440
Db 300 TTCCGAGGACGCGCAATCGAACAGATGGGGCTGGGCTGGGAGACGACTACGGCTCCG 241
Qy 1441 GAAACGCAATATACCATCACAGTGGCTGGAGAGCGCTGTGTCG 1487
Db 240 GGAAGCGCGCGATACCAATCACAGTGGTATCGAAGGCGCCCTGGAGC 194

RESULT 24

AW179968
LOCUS
DEFINITION
MGA0020f Mga Library Mycosphaerella graminicola cDNA clone MGA0020
5' similar to catalase-peroxidase, mRNA sequence.
AW179968
AW179968.1 GI:6447163
EST.
SOURCE
Mycosphaerella graminicola
Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyrionomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella
1 (bases 1 to 729)

Keon,J.P.R., Bailey,A.M. and Hargreaves,J.A.
A group of expressed cDNA sequences from the wheat fungal leaf
blotch pathogen, Mycosphaerella graminicola (Septoria tritici)
Fungal Genet. Biol. 29 (2), 118-133 (2000)
JOURNAL
MEDLINE
20374020
PUBMED
10919380
Contact: Hargreaves JA
Cell Biology Department
IACR-Long Ashton Research Station
Long Ashton, Bristol, BS41 9AF, UK

REFERENCE

AW179968
LOCUS
DEFINITION
MGA0020f Mga Library Mycosphaerella graminicola cDNA clone MGA0020
5' similar to catalase-peroxidase, mRNA sequence.
AW179968
AW179968.1 GI:6447163
EST.
SOURCE
Mycosphaerella graminicola
Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyrionomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella
1 (bases 1 to 729)

REFERENCE

AW179968
LOCUS
DEFINITION
MGA0020f Mga Library Mycosphaerella graminicola cDNA clone MGA0020
5' similar to catalase-peroxidase, mRNA sequence.
AW179968
AW179968.1 GI:6447163
EST.
SOURCE
Mycosphaerella graminicola
Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyrionomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella
1 (bases 1 to 729)

REFERENCE

AW179968
LOCUS
DEFINITION
MGA0020f Mga Library Mycosphaerella graminicola cDNA clone MGA0020
5' similar to catalase-peroxidase, mRNA sequence.
AW179968
AW179968.1 GI:6447163
EST.
SOURCE
Mycosphaerella graminicola
Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothyrionomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella
1 (bases 1 to 729)

Tel: +44(0)1275 392181
Fax: +44(0)1275 394281
Email: john.hargreaves@barc.ac.uk
Insert Length: 2600 Std Error: 0.00
Seq primer: M13 reverse.

FEATURES

source

1..729
/organism="Mycosphaerella graminicola"
/mol_type="mRNA"
/strain="Strit"
/db_xref="taxon:54734"
/clone="MGA0020"
/clone_lib="Mga Library"
/note="vector: pSPOR1; Library constructed from cultures
utilizing ammonium ions as a source of nitrogen"

ORIGIN

Query Match 11.9%; Score 176.8; DB 2; Length 729;
Best Local Similarity 61.0%; Pred. No. 1.3e-42;
Matches 305; Conservative 0; Mismatches 192; Indels 3; Gaps 1;
Qy 632 CTACTATCCAGAAACACTGGATTTAACTCTCTGAGATTACACAGCCCTGAATCAATCC 691
Db 109 CTGTGGCCAAATGAGCTGAACACCAAGATCTTCGCGCAGCACACCGCTGCTACCGACC 168
Qy 692 CTGGGGGGCTGATTTGATTTATGCCACAGATTTCAACAGCTGATATGAGAGCTCTGAA 751
Db 169 ATTGGGCAAGCAGTTGACTACCCAGCCGATTTCAAGAGCTTTGATTTACAAATGGCCTCA 228
Qy 752 AAAAGATATCAAGATTTGCTGACAACTTCCAGAGTTGGTCCCTCGGATTTATGCTCA 811
Db 229 GAAGGACTTGAACGACTCATGACCGACTCGAAGGACTTCTGGCCAGCAGACTTTGGCCA 288
Qy 812 TTATGTCCTTCTTTATTCGTATGGCTTGGCAGCGTGCAGCAATACATACAGCATATGA 871
Db 289 CTAGGTTGTTCTTCGTCGCGATGCGATGCGCAGCGCTGTTACCTACCGTTCGCTGA 348
Qy 872 TGCGCGGGAGCGCCAGTGGTTCAGCAACGTTTGAACCGCTGAACAGCTGGCGCGGA 931
Db 349 TGGTCTGCTGGTGGTGGTGGAGGACAGCAGCGATTCGCTCTCTCAACTCATGGCGGA 408
Qy 932 TAACTGTTAATCTGATAAAGCCCGTCGATTTGTCGCCAGTCAAGAAAAAATACGGCTC 991
Db 409 CAAGCGCAATCTCGAAGGCTCGCCGCTTGTGGCCAAATCAAGAAAAGTACGGAAA 468
Qy 992 CAGTATTTCTGGGAGACCTGATGTCCTGACTGTAATTTGTCCTTGAATCCATGGG 1051
Db 469 CAAGATCTCATGGCGGACCTGTTCTCTCTCACTGGAATGTGCGCAATTGAGTCTATGG 528
Qy 1052 ATTTAAACCTGGGATTTCTGCGGGAAGAGATGACTGGGAGTCG---GACCTGGT 1108
Db 529 TCTGCCACACGTTTGGTTTCGCGCGTGGCCGTCGACACATGGGAGCGGATGACTCCGT 588
Qy 1109 ATACTGGGCGCTGACAACA 1128
Db 589 ATACTGGGCTGGTGAACCA 608

RESULT 25

AW179968

LOCUS

DEFINITION

AW179968

AW179968.1

GI:13259728

EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

562 bp mRNA linear EST 09-MAR-2001
Lma123 v8 mycelial library Leptosphaeria maculans cDNA clone Lma123
5' similar to putative catalase/peroxidase, mRNA sequence.

Leptosphaeria maculans (blackleg of rapeseed fungus)
Leptosphaeria maculans
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
Pleosporales; Leptosphaeriaceae; Leptosphaeria; Leptosphaeria
maculans complex.

1 (bases 1 to 562)
Idnurm,A., Cozijnsen,A.J. and Howlett,B.J.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 39.9662 Seconds
(without alignments)
1190.710 Million cell updates/sec

Title: US-09-674-277-1_COPY_400_407

Perfect score: 8

Sequence: 1 atcgctcag 8

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	8	100.0	15	9	US-09-996-561-61
C 2	8	100.0	15	9	US-09-884-948-61
C 3	8	100.0	15	10	US-09-888-326-636
C 4	8	100.0	15	10	US-09-776-479-166
C 5	8	100.0	15	11	US-09-776-479-166
C 6	8	100.0	15	14	US-10-112-653-159
C 7	8	100.0	15	14	US-10-017-995-166
C 8	8	100.0	15	17	US-10-314-578-166
C 9	8	100.0	15	18	US-10-600-286-3
C 10	8	100.0	15	18	US-10-831-778-166
C 11	8	100.0	16	17	US-10-138-674-7009

C 12	8	100.0	16	17	US-10-138-674-7010	Sequence 7010, Ap
C 13	8	100.0	16	18	US-10-287-949A-7009	Sequence 7009, Ap
C 14	8	100.0	16	18	US-10-287-949A-7010	Sequence 7010, Ap
C 15	8	100.0	16	19	US-10-722-155-15	Sequence 15, Appl
C 16	8	100.0	17	9	US-09-866-108-8750	Sequence 8750, Ap
C 17	8	100.0	17	9	US-09-866-108-8751	Sequence 8751, Ap
C 18	8	100.0	17	9	US-09-866-108-8752	Sequence 8752, Ap
C 19	8	100.0	17	9	US-09-866-108-8753	Sequence 8753, Ap
C 20	8	100.0	17	9	US-09-866-108-8754	Sequence 8754, Ap
C 21	8	100.0	17	9	US-09-866-108-8755	Sequence 8755, Ap
C 22	8	100.0	17	9	US-09-866-108-8756	Sequence 8756, Ap
C 23	8	100.0	17	9	US-09-866-108-8757	Sequence 8757, Ap
C 24	8	100.0	17	9	US-09-866-108-8758	Sequence 8758, Ap
C 25	8	100.0	17	9	US-09-866-108-8759	Sequence 8759, Ap
C 26	8	100.0	17	10	US-09-846-903-29	Sequence 29, Appl
C 27	8	100.0	17	16	US-10-339-782-132	Sequence 132, Appl
C 28	8	100.0	17	16	US-10-150-045-10	Sequence 10, Appl
C 29	8	100.0	17	17	US-10-138-674-1699	Sequence 1699, Ap
C 30	8	100.0	17	17	US-10-138-674-1700	Sequence 1700, Ap
C 31	8	100.0	17	17	US-10-138-674-6326	Sequence 6326, Ap
C 32	8	100.0	17	17	US-10-138-674-8632	Sequence 8632, Ap
C 33	8	100.0	17	17	US-10-138-674-8633	Sequence 8633, Ap
C 34	8	100.0	17	18	US-10-287-949A-1699	Sequence 1699, Ap
C 35	8	100.0	17	18	US-10-287-949A-1700	Sequence 1700, Ap
C 36	8	100.0	17	18	US-10-287-949A-6326	Sequence 6326, Ap
C 37	8	100.0	17	18	US-10-287-949A-8632	Sequence 8632, Ap
C 38	8	100.0	17	18	US-10-287-949A-8633	Sequence 8633, Ap
C 39	8	100.0	17	18	US-10-660-208-29	Sequence 29, Appl
C 40	8	100.0	17	18	US-10-723-361-8750	Sequence 8750, Ap
C 41	8	100.0	17	18	US-10-723-361-8751	Sequence 8751, Ap
C 42	8	100.0	17	18	US-10-723-361-8752	Sequence 8752, Ap
C 43	8	100.0	17	18	US-10-723-361-8753	Sequence 8753, Ap
C 44	8	100.0	17	18	US-10-723-361-8754	Sequence 8754, Ap
C 45	8	100.0	17	18	US-10-723-361-8755	Sequence 8755, Ap

ALIGNMENTS

RESULT 1

US-09-996-561-61/c

; Sequence 61, Application US/09996561

; Patent No. US20020151703A1

; GENERAL INFORMATION:

; APPLICANT: YOKOYAMA, KEIICHI

; APPLICANT: NAKAMURA, NAMI

; APPLICANT: MIWA, TETSUYA

; APPLICANT: SEGURO, KATSUYA

; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE

; FILE REFERENCE: 0010-0937-0

; CURRENT APPLICATION NUMBER: US/09/996,561

; CURRENT FILING DATE: 2001-11-30

; PRIOR APPLICATION NUMBER: CURRENT APPLICATION NUMBER: US/09/448,310

; PRIOR FILING DATE: CURRENT FILING DATE: 1999-11-24

; PRIOR APPLICATION NUMBER: 09/109,063

; PRIOR FILING DATE: 1998-07-02

; NUMBER OF SEQ ID NOS: 62

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 61

; LENGTH: 15

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: CODON FOR

; OTHER INFORMATION: N-TERMINAL FRAGMENT

US-09-996-561-61

Query Match 100.0%; Score 8; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8

|||||||

```
Db          9  ATCGTCAG 2

RESULT 2
US-09-884-948-61/c
; Sequence 61, Application US/09884948
; Patent No. US20020173021A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, KEIICHI
; APPLICANT: NAKAMURA, NAMI
; APPLICANT: MIWA, TETSUYA
; APPLICANT: SEGURO, KATSUYA
; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE
; FILE REFERENCE: 0010-0937-0
; CURRENT APPLICATION NUMBER: US/09/884,948
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 09/448,310
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CODON FOR
; OTHER INFORMATION: N-TERMINAL FRAGMENT
US-09-884-948-61

Query Match      100.0%; Score 8; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05; Length 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1  ATCGTCAG 8
           |||||
Db          9  ATCGTCAG 2

RESULT 3
US-09-888-326-636/c
; Sequence 636, Application US/09888326
; Publication No. US20030026801A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Hartmann, Gunther
; TITLE OF INVENTION: Methods for Enhancing Antibody-Induced
; TITLE OF INVENTION: Cell Lysis and Treating Cancer
; FILE REFERENCE: C1039/7052 (AWS)
; CURRENT APPLICATION NUMBER: US/09/888,326
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,346
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 636
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: phosphodiester backbone
US-09-888-326-636

Query Match      100.0%; Score 8; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05; Length 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1  ATCGTCAG 8
           |||||
Db          10 ATCGTCAG 3

RESULT 4
US-09-776-479-166/c
; Sequence 166, Application US/09776479
; Publication No. US20030087848A1
; GENERAL INFORMATION:
; APPLICANT: Bratzler, Robert L.
; APPLICANT: Petersen, Deanna M.
; APPLICANT: Fouron, Yves
; TITLE OF INVENTION: Immunostimulatory Nucleic Acids for the
; TITLE OF INVENTION: Treatment of Asthma and Allergy
; FILE REFERENCE: C1037/7013 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/09/776,479
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/179,991
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-776-479-166

Query Match      100.0%; Score 8; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05; Length 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1  ATCGTCAG 8
           |||||
Db          10 ATCGTCAG 3

RESULT 5
US-09-776-479-166/c
; Sequence 166, Application US/09776479
; Publication No. US20040067902A9
; GENERAL INFORMATION:
; APPLICANT: Bratzler, Robert L.
; APPLICANT: Petersen, Deanna M.
; APPLICANT: Fouron, Yves
; TITLE OF INVENTION: Immunostimulatory Nucleic Acids for the
; TITLE OF INVENTION: Treatment of Asthma and Allergy
; FILE REFERENCE: C1037/7013 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/09/776,479
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/179,991
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-09-776-479-166

Query Match      100.0%; Score 8; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05; Length 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          1  ATCGTCAG 8
           |||||
Db          10 ATCGTCAG 3

RESULT 6
US-10-112-653-159/c
; Sequence 159, Application US/10112653
; Publication No. US20030050268A1
```

GENERAL INFORMATION:
; APPLICANT: Krieg, Arthur M.
; APPLICANT: Berg, Daniel J.
; TITLE OF INVENTION: IMMUNOSTIMULATORY NUCLEIC ACID FOR
; TITLE OF INVENTION: TREATMENT OF NON-ALLERGIC INFLAMMATORY DISEASES
; FILE REFERENCE: C01039/70060 (AWS)
; CURRENT APPLICATION NUMBER: US/10/112,653
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 60/279,642
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 1040
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-112-653-159

Query Match 100.0%; Score 8; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 7

US-10-017-995-166/c
; Sequence 166, Application US/10017995
; Publication No. US20030055014A1
; GENERAL INFORMATION:
; APPLICANT: Bratzler, Robert L.
; TITLE OF INVENTION: Inhibition of Angiogenesis by Nucleic Acids
; FILE REFERENCE: C1037/7025 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/10/017,995
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/255,534
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-017-995-166

Query Match 100.0%; Score 8; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 8

US-10-314-578-166/c
; Sequence 166, Application US/10314578
; Publication No. US20030212026A1
; GENERAL INFORMATION:
; APPLICANT: Krieg, Arthur M.
; APPLICANT: Schetter, Christian
; APPLICANT: Vollmer, Jorg
; TITLE OF INVENTION: Immunostimulatory Nucleic Acids
; FILE REFERENCE: C1039/7035 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/10/314,578
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/156,113

; PRIOR FILING DATE: 1999-09-25
; PRIOR APPLICATION NUMBER: US 60/156,135
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: US 60/227,436
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 1145
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-314-578-166

Query Match 100.0%; Score 8; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 3

RESULT 9

US-10-600-286-3/c
; Sequence 3, Application US/10600286
; Publication No. US20040219556A1
; GENERAL INFORMATION:
; APPLICANT: BAZAN, GUILLERMO C.
; APPLICANT: GAYLORD, BRENT S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTION AND ANALYSIS OF
; TITLE OF INVENTION: POLYNUCLEOTIDES USING LIGHT HARVESTING
; TITLE OF INVENTION: MULTICHROMOPHORES
; FILE REFERENCE: 701995-2001
; CURRENT APPLICATION NUMBER: US/10/600,286
; CURRENT FILING DATE: 2003-06-21
; PRIOR APPLICATION NUMBER: 60/406,266
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 3
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
US-10-600-286-3

Query Match 100.0%; Score 8; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2

RESULT 10

US-10-831-778-166/c
; Sequence 166, Application US/10831778
; Publication No. US20040235774A1
; GENERAL INFORMATION:
; APPLICANT: Bratzler, Robert L.
; APPLICANT: Petersen, Deanna M.
; APPLICANT: Fouron, Yves
; TITLE OF INVENTION: Immunostimulatory Nucleic Acids for the
; TITLE OF INVENTION: Treatment of Asthma and Allergy
; FILE REFERENCE: C1037/7013 (HCL/MAT)
; CURRENT APPLICATION NUMBER: US/10/831,778
; CURRENT FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/179,991

```
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-831-778-166

Query Match      100.0%; Score 8; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATCGTCAG 8
        |||||
Db      10 ATCGTCAG 3

RESULT 11
US-10-138-674-7009/c
; Sequence 7009, Application US/10138674
; Publication No. US20040077565A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138,674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7009
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-7009

Query Match      100.0%; Score 8; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATCGTCAG 8
        |||||
Db      16 ATCGTCAG 9

RESULT 12
US-10-138-674-7010/c
; Sequence 7010, Application US/10138674
; Publication No. US20040077565A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/138,674
; CURRENT FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7010
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-138-674-7010
```

US-10-138-674-7010

Query Match 100.0%; Score 8; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
 |||||
Db 13 ATCGTCAG 6

RESULT 13
US-10-287-949A-7009/c
; Sequence 7009, Application US/10287949A
; Publication No. US20040102389A1
; GENERAL INFORMATION: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/287,949A
; CURRENT FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7009
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-287-949A-7009

Query Match 100.0%; Score 8; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
 |||||
Db 16 ATCGTCAG 9

RESULT 14
US-10-287-949A-7010/c
; Sequence 7010, Application US/10287949A
; Publication No. US20040102389A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-N (400/049)
; CURRENT APPLICATION NUMBER: US/10/287,949A
; CURRENT FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 20822
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7010
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-287-949A-7010

Query Match 100.0%; Score 8; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
 |||||
Db 13 ATCGTCAG 6

RESULT 15

US-10-722-155-15/c
; Sequence 15, Application US/10722155
; Publication No. US20050019786A1
; GENERAL INFORMATION:
; APPLICANT: LEPROUST, ERIC M.
; APPLICANT: WOLBER, PAUL K.
; TITLE OF INVENTION: METHODS AND APPARATUS FOR PREPARING ARRAYS COMPRISING
; FILE REFERENCE: 10030511-1
; CURRENT APPLICATION NUMBER: US/10/722,155
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 10/266,474
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 09/350,969
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 15
; LENGTH: 16
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
US-10-722-155-15

Query Match 100.0%; Score 8; DB 19; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2
|||||

RESULT 16

US-09-866-108-8750
; Sequence 8750, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 8750
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-8750

Query Match 100.0%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 10 ATCGTCAG 17
|||||

RESULT 17

US-09-866-108-8751
; Sequence 8751, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: PCT/US01/00663

		Best Local Similarity 100.0%; Pred. No. 1.1e+05; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY		1 ATCGTCAG 8	
DB		8 ATCGTCAG 15	
RESULT 19			
US-09-866-108-8753		Sequence 8753, Application US/09866108	
		Patent No. US20020048800A1	
		GENERAL INFORMATION:	
		APPLICANT: GU, Yizhong	
		APPLICANT: JI, Yonggang	
		APPLICANT: PENN, Sharron G.	
		APPLICANT: HANZEL, David K.	
		APPLICANT: RANK, David R.	
		APPLICANT: CHEN, Wensheng	
		APPLICANT: SHANNON, Mark	
		TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE	
		FILE REFERENCE: AEOMICA-7	
		CURRENT APPLICATION NUMBER: US/09/866,108	
		CURRENT FILING DATE: 2001-05-25	
		PRIOR APPLICATION NUMBER: US 60/207,456	
		PRIOR FILING DATE: 2000-05-26	
		PRIOR APPLICATION NUMBER: GB 24263.6	
		PRIOR FILING DATE: 2000-10-04	
		PRIOR APPLICATION NUMBER: US 60/236,359	
		PRIOR FILING DATE: 2000-09-27	
		PRIOR APPLICATION NUMBER: PCT/US01/00666	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00667	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00664	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00669	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00665	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00668	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00663	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00662	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00661	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00670	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: US 60/234,687	
		PRIOR FILING DATE: 2000-09-21	
		PRIOR APPLICATION NUMBER: US 60/266,860	
		PRIOR FILING DATE: 2001-02-05	
		NUMBER OF SEQ ID NOS: 15752	
		SOFTWARE: Aeomica Sequence Listing Engine	
		SEQ ID NO 8753	
		LENGTH: 17	
		TYPE: DNA	
		ORGANISM: Homo sapiens	
		US-09-866-108-8753	
Query Match		100.0%; Score 8; DB 9; Length 17;	
Best Local Similarity		100.0%; Pred. No. 1.1e+05; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY		1 ATCGTCAG 8	
DB		7 ATCGTCAG 14	
RESULT 20			
US-09-866-108-8754			

		Best Local Similarity 100.0%; Pred. No. 1.1e+05; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY		1 ATCGTCAG 8	
DB		9 ATCGTCAG 16	
RESULT 18			
US-09-866-108-8752		Sequence 8752, Application US/09866108	
		Patent No. US20020048800A1	
		GENERAL INFORMATION:	
		APPLICANT: GU, Yizhong	
		APPLICANT: JI, Yonggang	
		APPLICANT: PENN, Sharron G.	
		APPLICANT: HANZEL, David K.	
		APPLICANT: RANK, David R.	
		APPLICANT: CHEN, Wensheng	
		APPLICANT: SHANNON, Mark	
		TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE	
		FILE REFERENCE: AEOMICA-7	
		CURRENT APPLICATION NUMBER: US/09/866,108	
		CURRENT FILING DATE: 2001-05-25	
		PRIOR APPLICATION NUMBER: US 60/207,456	
		PRIOR FILING DATE: 2000-05-26	
		PRIOR APPLICATION NUMBER: GB 24263.6	
		PRIOR FILING DATE: 2000-10-04	
		PRIOR APPLICATION NUMBER: US 60/236,359	
		PRIOR FILING DATE: 2000-09-27	
		PRIOR APPLICATION NUMBER: PCT/US01/00666	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00667	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00664	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00669	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00665	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00668	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00663	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00662	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00661	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: PCT/US01/00670	
		PRIOR FILING DATE: 2001-01-30	
		PRIOR APPLICATION NUMBER: US 60/234,687	
		PRIOR FILING DATE: 2000-09-21	
		PRIOR APPLICATION NUMBER: US 60/266,860	
		PRIOR FILING DATE: 2001-02-05	
		NUMBER OF SEQ ID NOS: 15752	
		SOFTWARE: Aeomica Sequence Listing Engine	
		SEQ ID NO 8752	
		LENGTH: 17	
		TYPE: DNA	
		ORGANISM: Homo sapiens	
		US-09-866-108-8752	
Query Match		100.0%; Score 8; DB 9; Length 17;	

RESULT 22
US-09-866-108-8756
Sequence 8756, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharron G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wenaheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27

RESULT 22
US-09-866-108-8756
Sequence 8756, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharon G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEWICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/266,860
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 15752
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 8757
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapiens
US-09-866-108-8757

Query Match 100.0%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
| | | | | | | |
Db 3 ATCGTCAG 10

RESULT 24
US-09-866-108-8758
Sequence 8758, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharron G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/266,860
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 15752
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 8756
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapiens
US-09-866-108-8756

Query Match 100.0%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
| | | | | | | |
Db 4 ATCGTCAG 11

RESULT 23
US-09-866-108-8757
Sequence 8757, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharron G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 60/266,860
;; PRIOR FILING DATE: 2001-02-05
;; NUMBER OF SEQ ID NOS: 15752
;; SOFTWARE: Aecomica Sequence Listing Engine
;; SEQ ID NO 8758
;; LENGTH: 17
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-866-108-8758

Query Match 100.0%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 2 ATCGTCAG 9

RESULT 25
US-09-866-108-8759
; Sequence 8759, Application US/09866108
; Patent No. US2002004800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AECOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866.108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 8759
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-866-108-8759

Query Match 100.0%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCGTCAG 8
Db 1 ATCGTCAG 8

Search completed: March 12, 2005, 00:25:13
Job time : 39.9662 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 105.861 Seconds
(without alignments)
2876.537 Million cell updates/sec

Title: US-09-674-277-1_COPY_400_407

Perfect score: 8

Sequence: 1 atcgctcag 8

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	100.0	26	AZ313876	AZ313876 1M0030H21
2	8	100.0	29	BZ665244	BZ665244 SALK_1109
3	8	100.0	29	AG202191	AG202191 Pan trogl
4	8	100.0	32	BX891489	BX891489 Arabidops
5	8	100.0	34	AI366127	AI366127 ac94f04.x
6	8	100.0	34	AZ309524	AZ309524 1M0013104
7	8	100.0	35	AV954831	AV954831 AV954831
8	8	100.0	36	CR399031	CR399031 Arabidops
9	8	100.0	36	TA43F04Q	TA43F04Q
10	8	100.0	37	CF297401	CF297401 30DGS--08
11	8	100.0	37	AL769050	AL769050 Arabidops
12	8	100.0	38	CD745197	CD745197 RB3_E11.R
13	8	100.0	41	BH908980	BH908980 SALK_0515
14	8	100.0	46	BZ382637	BZ382637 SALK_1185
15	8	100.0	46	CR770256	CR770256 Arabidops
16	8	100.0	48	AA402658	AA402658 zu49e08.s
17	8	100.0	49	AI005836	AI005836 ua83f11.r
18	8	100.0	51	AL754099	AL754099 Arabidops
19	8	100.0	53	CC940957	CC940957 01S0630-0
20	8	100.0	54	CO168073	CO168073 Mdfir3005d
21	8	100.0	55	BX555677	BX555677 BX555677
22	8	100.0	55	T73110	T73110 yc67d06.s1
23	8	100.0	55	CR111119	CR111119 Forward s
24	8	100.0	55	DM546537	DM546537 Drosophil

C 25	8	100.0	55	9	CL610270	CL610270	EY10126-3
C 26	8	100.0	58	9	CG733518	CG733518	111915681
C 27	8	100.0	59	8	BZ763114	BZ763114	SALK_1134
C 28	8	100.0	59	9	CR109453	CR109453	Forward s
C 29	8	100.0	61	2	AW059948	AW059948	LF8a09.YG
C 30	8	100.0	61	5	BX621186	BX621186	BX621186
C 31	8	100.0	61	7	N93964	N93964	2b74g02.s1
C 32	8	100.0	61	9	BX893635	BX893635	Arabidops
C 33	8	100.0	61	9	CG733166	CG733166	1119154E1
C 34	8	100.0	61	9	CL234350	CL234350	0250422-0
C 35	8	100.0	62	1	AL774523	AL774523	AL774523
C 36	8	100.0	62	8	BH813130	BH813130	SALK_0637
C 37	8	100.0	62	8	BH814725	BH814725	SALK_0668
C 38	8	100.0	63	6	CD524103	CD524103	ku96B04.Y
C 39	8	100.0	63	8	BH812721	BH812721	SALK_0629
C 40	8	100.0	63	8	BZ383684	BZ383684	SALK_1342
C 41	8	100.0	63	8	CC456855	CC456855	SALK_1010
C 42	8	100.0	63	9	AJ595143	AJ595143	Arabidops
C 43	8	100.0	64	1	AA714573	AA714573	nx83e12.s
C 44	8	100.0	64	1	AA937456	AA937456	on40b07.s
C 45	8	100.0	64	5	BX558244	BX558244	BX558244

ALIGNMENTS

RESULT 1
AZ313876
LOCUS
DEFINITION
1M0030H21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0030H21 F, genomic survey sequence.

ACCESSION
AZ313876
VERSION
GSS.

KEYWORDS
GSS.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

REFERENCE
1 (bases 1 to 26)

AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0030 row: H column: 21
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 26.

FEATURES
Location/Qualifiers
1..26

source
/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0030H21"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 [gi4732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 8; DB 8; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 8 ATCGTCAG 15

RESULT 2

BZ665244/c 29 bp DNA linear GSS 31-JAN-2003
LOCUS
DEFINITION
Arabidopsis thaliana genomic clone SALK_110927.40.25.x, genomic survey sequence.

ACCESSION BZ665244.1 GI:28182689
VERSION
KEYWORDS

SOURCE

ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 29)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jenke,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

TITLE

JOURNAL

COMMENT
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At1g55870.
Class: TDNA tagged.
FEATURES
source
1. .29
/organism="Arabidopsis thaliana"
/mol type="genomic DNA"
/ecotype="Col-0"
/db xref="taxon:3702"
/clone="SALK_110927.40.25.x"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 100.0%; Score 8; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 13 ATCGTCAG 6

RESULT 3

AG202191/c 29 bp DNA linear GSS 06-MAR-2004
LOCUS
DEFINITION
Pan troglodytes DNA, clone: RP43-085D23.T7, genomic survey sequence.

ACCESSION AG202191.1 GI:45234366
VERSION
KEYWORDS

SOURCE

ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

AUTHORS
Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
BAC end sequences of Library RP-43

TITLE

JOURNAL

REFERENCE

AUTHORS
Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.

TITLE

JOURNAL

COMMENT
Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC); 52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
E-mail:redstone@mail.krribb.re.kr, URL:<http://phs.grc.krribb.re.kr/>,
Tel:82-42-866-7181, Fax:82-42-860-4409

Clones are derived from the chimpanzee BAC library RP-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

FEATURES

source

1. .29
/organism="Pan troglodytes"
/mol type="genomic DNA"
/db xref="taxon:9598"
/clone="RP43-085D23.T7"
/sex="male"
/cell type="lymphocytes"
/clone_lib="RP-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 100.0%; Score 8; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 27 ATCGTCAG 20

RESULT 4

BX891489/c 32 bp DNA linear GSS 05-APR-2004
LOCUS
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence GK-550A04-023856, genomic survey sequence.

ACCESSION BX891489.1 GI:39923984
VERSION
KEYWORDS

SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1
Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P., and Weisshaar, B.
GABI-Kat SimpleSearch: a flanking sequence tag (PST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL MEDLINE PUBLISHED 22755829
PUBMED 12874060

REFERENCE 2
Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
Plant Mol. Biol. 53 (1-2), 247-259 (2003)

JOURNAL MEDLINE PUBLISHED 23117147
PUBMED 14756321

REFERENCE 3
Strizhov, N., Li, Y., Rosso, M.G., Viehoever, P., Dekker, K.A. and Weisshaar, B.
High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
BioTechniques 35 (6), 1164-1168 (2003)

JOURNAL PUBLISHED 14682050

REFERENCE 4 (bases 1 to 32)
Li, Y., Rosso, M.G., Strizhov, N. and Weisshaar, B.
Direct Submission
Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene Atlg79450. Details on the protocols used for generations of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES
source
1..32
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-550A04-023856"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN
Query Match 100.0%; Score 8; DB 9; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||
Db 29 ATCGTCAG 22

RESULT 5
AI366127/c
LOCUS
ac94f04.x1 Schiller meningioma Homo sapiens cDNA clone
DEFINITION
IMAGE:1953535.3' similar to TR:Q14829 Q14829 MG61 ;contains element MSR1 repetitive element ;, mRNA sequence.

ACCESSION AI366127
VERSION AI366127.1 GI:4125816
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 34)
Hallier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)

TITLE
JOURNAL COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1..34
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1953535"
/sex="female"
/tissue_type="meningioma"
/dev_stage="72 years"
/lab_host="SOLR"
/clone_lib="Schiller meningioma"
/note="Organ: brain; Vector: pBluescript SK- (Stratagene); Site 1: EcoRI; Site 2: XhoI; Double-stranded cDNA was prepared from human meningioma using primer 5'-GAGAGAGAGAGAGAGAGAACTAGTCTCAGT(18)-3'. An EcoRI adaptor was used on the 5' end of the cDNA as follows: 5'-AATCGCAGCAG-3'. The library was size-selected and went through one round of amplification. Average insert size is 1.7 kb, with a range from 0.4-12 kb. Tumor identification by consensus pathology. This library was constructed by Dr. Martin Schiller (Johns Hopkins University)."

ORIGIN
Query Match 100.0%; Score 8; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||
Db 16 ATCGTCAG 9

RESULT 6
AZ309524/c
LOCUS
DEFINITION
AZ309524 Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0013104 R, genomic survey sequence.

ACCESSION AZ309524.1 GI:10350592
VERSION
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 34)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddm@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0013 row: I column: 04
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 34.

FEATURES
 source
 1. .34
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCIM0013104"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGCIM library"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
 Query Match 100.0%; Score 8; DB 8; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 13 ATCGTCAG 6

RESULT 7
 AV954831/c
 LOCUS
 DEFINITION
 Ciona intestinalis cDNA library, cleavage stage embryo
 AV954831
 VERSION
 AV954831.1 GI:19443130
 KEYWORDS
 EST.
 SOURCE
 Ciona intestinalis
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.
 REFERENCE
 1 (bases 1 to 35)
 AUTHORS
 Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.

Expressed genes in Ciona intestinalis
 Unpublished (2000)
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
 source
 1. .35
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cic108f20"
 /tissue_type="whole animal"
 /dev_stage="cleavage stage embryo"
 /clone_lib="Nori Satoh unpublished cDNA library, cleavage stage embryo"

ORIGIN
 Query Match 100.0%; Score 8; DB 2; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
 |||||
 Db 34 ATCGTCAG 27

RESULT 8
 CR399031
 LOCUS
 DEFINITION
 Arabidopsis thaliana T-DNA flanking sequence GK-814D04-025642, genomic survey sequence.
 CR399031
 VERSION
 CR399031.1 GI:46939759
 KEYWORDS
 GSS.
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
 1
 Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weisshaar, B.
 GABI-Kat SimpleSearch: a flanking sequence tag (PST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
 Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL
 MEDLINE
 PUBMED
 22755829
 12874060

REFERENCE
 2
 Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B.
 An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
 Plant Mol. Biol. 53 (1-2), 247-259 (2003)

JOURNAL
 MEDLINE
 PUBMED
 23117147
 14756321

REFERENCE
 3
 Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weisshaar, B.
 High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
 Biotechniques 35 (6), 1164-1168 (2003)

JOURNAL
 PUBMED
 14682050

REFERENCE
 4 (bases 1 to 36)
 Strizhov, N., Li, Y., Rosso, M.G. and Weisshaar, B.
 Direct Submission
 Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
 This sequence has been recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by BAC clone

T7p1. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

Location/Qualifiers
1..36
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-814D04-025642"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"

/notes="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC106 (GenBank accession number: AJ537513). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 100.0%; Score 8; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||||
Db 27 ATCGTCAG 34

RESULT 9

TA43F04Q/c
LOCUS
DEFINITION
36 bp DNA linear GSS 13-DEC-2000
T. brucei sheared genomic DNA clone 43f04, reverse sequence,
genomic survey sequence.

ACCESSION
AL455253
VERSION
AL455253.1 GI:11856651
KEYWORDS
GSS.

SOURCE
Trypanosoma brucei

ORGANISM
Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.

REFERENCE
1 (bases 1 to 36)

AUTHORS
Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
Melville, S.E., Rajandream, M.A. and Barrell, B.G.

TITLE
Direct Submission

JOURNAL
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrel@sanger.ac.uk and
nhl@sanger.ac.uk

COMMENT
Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

Location/Qualifiers
1..36
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="43f04"

ORIGIN

Query Match 100.0%; Score 8; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||||
Db 15 ATCGTCAG 8

RESULT 10

CF297401/c
LOCUS
DEFINITION
37 bp mRNA linear EST 14-AUG-2003
30DGS--08-D24.g1 Rice leaf plasmid cDNA library I (30DGS) Oryza
sativa (japonica cultivar-group) cDNA clone 30DGS--08-D24, mRNA
sequence.

ACCESSION
CF297401
VERSION
CF297401.1 GI:336666434
KEYWORDS
EST.

SOURCE
Oryza sativa (japonica cultivar-group)

ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
1 (bases 1 to 37)

AUTHORS
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

TITLE
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)

JOURNAL
Contact: Nahm B.H.

COMMENT
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355

Email: bnahm@gbio.com, bnahm@bio.myongji.ac.kr.

Location/Qualifiers

1..37

FEATURES

Source
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="30DGS--08-D24"
/tissue_type="leaf"
/dev_stage="30 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library I (30DGS)"
/notes="Vector: PCR4-TORO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN

Query Match 100.0%; Score 8; DB 7; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||||
Db 13 ATCGTCAG 6

RESULT 11

AL769050
LOCUS
DEFINITION
37 bp DNA linear GSS 01-APR-2004
Arabidopsis thaliana T-DNA flanking sequence GK-083A06-011868,
genomic survey sequence.

ACCESSION
AL769050
VERSION
AL769050.1 GI:21522169
KEYWORDS
GSS.

SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1
REFERENCE
AUTHORS
TITLE
 Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weisshaar, B.
 GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
JOURNAL
MEDLINE
PUBMED
 Bioinformatics 19 (11), 1441-1442 (2003)
 23755829
 12874060

2
REFERENCE
AUTHORS
 Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B.
TITLE
 An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
JOURNAL
MEDLINE
PUBMED
 Plant Mol. Biol. 53 (1-2), 247-259 (2003)
 23117147
 14756321

3
REFERENCE
AUTHORS
 Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weisshaar, B.
TITLE
 High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
JOURNAL
MEDLINE
PUBMED
 BioTechniques 35 (6), 1164-1168 (2003)
 14682050

4 (bases 1 to 37)
REFERENCE
AUTHORS
TITLE
JOURNAL
 Li, Y., Rosso, M.G., Strizhov, N. and Weisshaar, B.
 Direct Submission
 Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT
 This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At4g26580. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES
 source
 1..37
 /location=Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Col-0"
 /notes="PCR was performed on DNA from Arabidopsis thaliana plants (TI) which were transformed with the T-DNA from vector pAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN
 Query Match 100.0%; Score 8; DB 9; Length 37;
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
 |||||
 Db 20 ATCGTCAG 27

RESULT 12
CD745197/c
LOCUS
DEFINITION
 RB3_E11_RB3_084 Naive Rat Blood-fed (RB) An.gam. 30 hr Abdomen Library Anopheles gambiae CDNA 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
 CD745197.1 GI:49249154
 EST.

SOURCE
ORGANISM
 Anopheles gambiae (African malaria mosquito)
 Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera, Culicoidea; Anopheles.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
 1 (bases 1 to 38)
 Dana, A.N., Lobo, N.F., Hillenmeyer, M.E. and Collins, F.H.
 Hematophagy-associated gene expression patterns in adult female Anopheles gambiae mosquitoes
 Unpublished (2003)
 Contact: Dana A.N.
 Frank H. Collins Laboratory
 University of Notre Dame
 Center for Tropical Disease Research and Training, Dept. of Biol.
 Sci., Notre Dame, IN 46556, USA
 Tel: 574 - 631 - 3241
 Fax: 574 - 631 - 3996
 Email: adana@nd.edu
 PCR Primers
 FORWARD: ctgggaagcgcgcattgtgtg
 BACKWARD: atagactcactataggcgaattggc
 Seq primer: ctgggaagcgcgcattgtgtg.
 Location/Qualifiers
 1..38
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="4Arr"
 /db_xref="taxon:7165"
 /sex="female"
 /issue_type="Abdomens"
 /dev_stage="Female adult 5-7 days post eclosion"
 /lab_host="E. coli XLI-Blue"
 /clone_lib="Naive Rat Blood-fed (RB) An.gam. 30 hr Abdomen Library"
 /notes="Vector: lambdaTriplEx2 (Clontech); Site_1: Sfi IA; Site 2: Sfi IB; Naive Rat Blood-fed adult female An. gambiae mosquitoes were flash frozen after a 30 hour incubation of adult mosquitoes at 19 degrees Celsius. Total RNA extracted from abdomens separated from remaining carcasses. CDNA inserts >500 bp cloned directionally into lTriplEx2; Sfi IA site is 5'. Non-normalized and Non-amplified phagemid library. Single pass sequencing reactions from 5' end."

ORIGIN
 Query Match 100.0%; Score 8; DB 6; Length 38;
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
 |||||
 Db 30 ATCGTCAG 23

RESULT 13
BH908980
LOCUS
DEFINITION
 SALK_051598.21.55.x Arabidopsis thaliana T-DNA insertion lines Arabidopsis thaliana genomic clone SALK_051598.21.55.x, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 41)
 Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J. and Ecker, J.R.
 A Sequence-Indexed Library of Insertion Mutations in the

JOURNAL COMMENT

Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
The Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.
Location/Qualifiers
1. .41
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clones="SALK_051598.21.55.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 100.0%; Score 8; DB 8; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||
Db 5 ATCGTCAG 12

RESULT 14
BZ382637/c
LOCUS
DEFINITION
SALK_118581.18.00.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_118581.18.00.x, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.
Location/Qualifiers
1. .46
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"

FEATURES
source

Query Match 100.0%; Score 8; DB 8; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||
Db 5 ATCGTCAG 12

RESULT 14
BZ382637/c
LOCUS
DEFINITION
SALK_118581.18.00.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_118581.18.00.x, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.

Class: TDNA tagged.
Location/Qualifiers
1. .46
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"

FEATURES
source

Query Match 100.0%; Score 8; DB 8; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||
Db 14 ATCGTCAG 7

RESULT 15
CR770256
LOCUS
DEFINITION
CR770256
Arabidopsis thaliana T-DNA flanking sequence GK-150H03-028303, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P., and Weissshaar, B.
GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE
22755829
PUBMED
12874060

REFERENCE
AUTHORS
TITLE

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weissshaar, B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE
23117147
PUBMED
14756321

REFERENCE
AUTHORS
TITLE

Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weissshaar, B.
High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
Biotechniques 35 (6), 1164-1168 (2003)
MEDLINE
23044198
PUBMED
14682050

REFERENCE
AUTHORS
TITLE

Rosso, M.G., Strizhov, N., Li, Y. and Weissshaar, B.
Direct Submission
Submitted (15-SEP-2004) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g37475.
Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

Location/Qualifiers
1. .46
/organism="Arabidopsis thaliana"

FEATURES
source

/db_xref="taxon:3702"
/clone="SALK_118581.18.00.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 100.0%; Score 8; DB 8; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
|||||

Db 14 ATCGTCAG 7

RESULT 15

CR770256

LOCUS

DEFINITION

CR770256

Arabidopsis thaliana T-DNA flanking sequence GK-150H03-028303, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weissshaar, B.

GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana

Bioinformatics 19 (11), 1441-1442 (2003)

MEDLINE

22755829

PUBMED

12874060

REFERENCE

AUTHORS

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weissshaar, B.

An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics

Plant Mol. Biol. 53 (1-2), 247-259 (2003)

MEDLINE

23117147

PUBMED

14756321

REFERENCE

AUTHORS

Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weissshaar, B.

High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines

Biotechniques 35 (6), 1164-1168 (2003)

MEDLINE

23044198

PUBMED

14682050

REFERENCE

AUTHORS

Rosso, M.G., Strizhov, N., Li, Y. and Weissshaar, B.

Direct Submission

Submitted (15-SEP-2004) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g37475.

Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

Location/Qualifiers

1. .46

/organism="Arabidopsis thaliana"

FEATURES

source

/mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="GK-150H03-028303"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /scot_type="Col-0"
 /notes="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161 (Genbank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 100.0%; Score 8; DB 9; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
 |||||
 Db 38 ATCGTCAG 45

RESULT 16

AA402658/c
 LOCUS
 DEFINITION
 z449e08.at Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:741350 3', similar to gb:X69181 60S RIBOSOMAL PROTEIN L31 (HUMAN); mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AA402658.1 GI:2056415
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 48)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 WashU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Insert Length: 1170 Std Error: 0.00
 Seq primer: -41m13 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source
 1. 48
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:5941428"
 /db_xref="taxon:9606"
 /clone="IMAGE:741350"
 /sex="Female"
 /tissue_type="ovarian tumor"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares ovary tumor NbHOT"
 /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into

ORIGIN

Query Match 100.0%; Score 8; DB 1; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATCGTCAG 8
 |||||
 Db 10 ATCGTCAG 3

RESULT 17

AI005836
 LOCUS
 DEFINITION
 ua83f11.r1 Soares mammary gland NbMMG Mus musculus cDNA clone IMAGE:1364109 5', similar to TR:P79251 P79251 VACUOLAR H(+)-ATPASE SUBUNIT.; mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AI005836.1 GI:3215445
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 49)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:897329
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source
 1. 49
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1364109"
 /sex="male"
 /tissue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares mammary gland NbMMG"
 /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."

ORIGIN


```

VERSION      CO168073.1  GI:48938611
KEYWORDS     EST.
SOURCE       Malus x domestica (cultivated apple)
ORGANISM     Malus x domestica
REFERENCE    1 (bases 1 to 54)
AUTHORS      Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A.,
              Aldwinckle, H., Mainoy, M., Carroll, N., Goldsbrough, P., Orvis, K.,
              Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T.,
              Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I.,
              Tsagaris, V., Kennedy, S., Waterston, R., and Wilson, R.
              Apple Functional Genomics grant - NSF 0321702
              Unpublished (2004)
TITLE        Contact: Schuyler S. Korban
JOURNAL      Apple Functional Genomics grant - NSF 0321702
COMMENT      Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              Library materials provided by: Schuyler S. Korban Library
              constructed by: D.O.Gonzalez/L. Vodkin Library sequenced by:
              Washington University Genome Sequencing Center
              WashU EST name: ash85b12.y1
              Seq primer: -40RP from Gibco
              High quality sequence stop: 54.
              Location/Qualifiers
FEATURES     source
             1..54
                /organism="Malus x domestica"
                /mol_type="mRNA"
                /db_xref="taxon:3750"
                /clone="Wdfr3005d23"
                /lab_host="DH10B ampicillin resistant"
                /clone_lib="Wdfr"
                /notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; Total
              RNA was extracted from freeze dried fruit tissue Stage 1-
              Young Fruitlet < 1 cm, using the method described by Wang
              and Vodkin (Plant Molecular Biology Reporter 12:132-145,
              1994). Poly(A)+mRNA was isolated from total RNA using the
              PolyAtract mRNA isolation system III (Promega). The
              library was prepared using the Invitrogen Life
              Technologies, Superscript Plasmid System with gateway
              technology for cDNA synthesis and cloning. Complementary
              DNA was synthesized from mRNA using a Poly (dT) sequence
              with a NotI restriction site. Sal I linker adapters were
              ligated to the blunt ended cDNA fragments followed by
              restriction with NotI. The cDNA fragments were
              directionally cloned into the NotI-SalI restriction site
              of the pSPORT 1 vector. The ligated cDNA fragments were
              transformed into E. coli Electromax DH10B host cells.
              Transformation efficiency: 10E+06 colony forming units
              Average insert size by PCR: 992 bp"
ORIGIN
             100.0%; Score 8; DB 7; Length 54;
             Best Local Similarity 100.0%; Pred. No. 1.7e+05;
             Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 15 ATCGTCAG 22

RESULT 21
BX555677/c BX555677 Glossina morsitans morsitans adult infected gut Glossina
LOCUS      BX555677 55 bp mRNA linear EST 10-OCT-2003
DEFINITION morsitans morsitans cDNA clone Tse22b07_glc, mRNA sequence.
ACCESSION BX555677
VERSION    BX555677
KEYWORDS   EST.
ORGANISM   Glossina morsitans morsitans
REFERENCE  1 (bases 1 to 55)
AUTHORS     Lehane, M.J., Aksoy, S., Gibson, W., Kerhornou, A., Berriman, M.,
            Hamilton, J., Soares, M.B., Bonaldo, M.F., Lehane, S. and Hall, N.
            Adult midgut expressed sequence tags from the tsetse fly Glossina
            morsitans morsitans and expression analysis of putative immune
            response genes
            Genome Biol. 4 (10), R63 (2003)
            MEDLINE 22881942
            PUBMED 14519198
            COMMENT Contact: Hall N
                    Pathogen Sequencing Unit
                    The Sanger Institute The Wellcome Trust Genome Campus
                    Hinxton, Cambridge, CB10 1SA, UK
                    Request for clones, please contact: Mike Lehane
                    Prof. M.J.Lehane
                    School of Biological Sciences,
                    University of Wales,
                    Bangor LL57 2UW
                    All clones with suffix glc are reverse primer reads starting at 5'
                    end of the cDNA all plc reads are from
                    the 3' end.
                    Location/Qualifiers
FEATURES     source
             1..55
                /organism="Glossina morsitans morsitans"
                /mol_type="mRNA"
                /sub_species="morsitans"
                /db_xref="taxon:37546"
                /clone="Tse22b07_glc"
                /issue_type="adult infected gut"
                /clone_lib="Glossina morsitans morsitans adult infected
              gut"
                /notes="country: Zimbabwe; EST from adult gut infected with
              T.brucei"
ORIGIN
             100.0%; Score 8; DB 5; Length 55;
             Best Local Similarity 100.0%; Pred. No. 1.7e+05;
             Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 53 ATCGTCAG 46

RESULT 22
T73110
LOCUS      T73110 55 bp mRNA linear EST 02-MAR-1995
DEFINITION yc67d06.s1 Stragatene liver (#937224) Homo sapiens cDNA clone
            IMAGE:85739 3', similar to gp:A06977 SERUM ALBUMIN PRECURSOR
            (HUMAN); mRNA sequence.
ACCESSION T73110
VERSION    T73110.1 GI:689785
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 55)
AUTHORS     Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
            Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
            Hawkins, M., Huitman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
            Hawking, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
            Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
            Travaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.,
            and Warra, M.
            Generation and analysis of 280,000 human expressed sequence tags
            Genome Res. 6 (9), 807-828 (1996)
            MEDLINE 97044478

```


REFERENCE
AUTHORS Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 55)
Levis, R., Hoskins, R., Liao, G., Mozdén, N., Tsang, G., He, Y.,
Karpen, G., Bellen, H., Rubin, G. and Spradling, A.
TITLE The Berkeley Drosophila Genome Project Gene Disruption Project
JOURNAL Unpublished (2001)
COMMENT Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106433947
Email: gerry@fruitfly.berkeley.edu
Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P
element

The P element insertion position is base 1 in the 55 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.
Class: transposon-tagged.

FEATURES
source
1..55
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster P{EPgy2} P element
insertion lines"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains one or more
P{EPgy2} P-element transposon insertion. The resultant
fragment for each strain was directly sequenced to
determine the genomic sequence at the site of insertion.
Details of the protocols used can be found at
<http://www.fruitfly.org/about/methods/inverse.pcr.html>."

ORIGIN
Query Match 100.0%; Score 8; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ATCGTCAG 8
Db 41 ATCGTCAG 34
|||||||

Search completed: March 11, 2005, 13:01:02
Job time : 109.861 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 2178.3 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-1

Perfect score: 1489

Sequence: 1 ctgcagtcgagatgaag.....ctggaaggaccttggtcgac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseqn_16Dec04:*

- 1: Geneseqn1980a:*
- 2: Geneseqn1980b:*
- 3: Geneseqn2000a:*
- 4: Geneseqn2001a:*
- 5: Geneseqn2001b:*
- 6: Geneseqn2002a:*
- 7: Geneseqn2002b:*
- 8: Geneseqn2003a:*
- 9: Geneseqn2003b:*
- 10: Geneseqn2003c:*
- 11: Geneseqn2003d:*
- 12: Geneseqn2004a:*
- 13: Geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1489	100.0	1489	3	Aaz36101 Nucleic a
2	521	35.0	2214	8	Aca54480 Prokaryot
3	394.6	26.5	2205	13	Adt44922 Bacterial
4	382	25.7	1358	8	Aca37119 Prokaryot
5	375	25.2	2139	13	AdS58848 Bacterial
6	357.6	24.0	2238	2	Aav06555 Microscil
7	357.6	24.0	2238	6	AbS53945 DNA encod
8	344.8	23.2	2145	13	AdS57655 Bacterial
9	334.6	22.5	2208	13	Adt42409 Bacterial
10	330.6	22.2	2163	13	AdS56510 Bacterial
11	323.2	21.7	2271	13	AdS60256 Bacterial
12	321.4	21.6	2223	12	Adh12940 Francisel
13	310.4	20.8	2205	13	Adt45831 Bacterial
14	305.4	20.5	2181	8	Aca19012 Prokaryot
15	305.4	20.5	2181	13	Adt48892 Bacterial
16	303.4	20.4	2166	13	AdS60748 Bacterial
17	302.2	20.3	2214	13	Adt42268 Bacterial
18	298.4	20.0	978	8	Aca38594 Prokaryot
19	298.4	20.0	2223	8	Aca40570 Prokaryot
20	298.4	20.0	2235	2	Adt89967 Mycobacte

21	298.4	20.0	2331	2	AAT90400	Mycobacte
22	298.4	20.0	2331	2	AAX00817	M. tuberc
c 23	298.4	20.0	82993	6	ABX09140	Mycobacte
c 24	298.4	20.0	110000	4	AAI99682	Continuati
c 25	298.4	20.0	110000	4	AAI99683	Continuati
26	297.6	20.0	2133	13	ADT43224	Bacterial
27	297.2	20.0	2226	13	ADT46662	Bacterial
28	296.8	19.9	2331	2	AAX00818	M. tuberc
29	293.6	19.7	2244	8	ACA26056	Prokaryot
30	289.8	19.5	2181	8	ACA51023	Prokaryot
31	288.2	19.4	2181	8	ACA51893	Prokaryot
32	286.8	19.3	2265	13	ADs47978	Bacterial
33	285.8	19.2	2175	8	ACA53244	Prokaryot
c 34	284.8	19.1	1674	11	ACH96706	Klebsiell
35	284.8	19.1	2175	8	ACA36062	Prokaryot
36	284.8	19.1	2208	11	ACH96581	Klebsiell
37	279.2	18.8	2094	8	ACA37985	Prokaryot
38	276.8	18.6	2144	8	ACA32201	Prokaryot
39	274.2	18.4	2221	2	AAT89968	Mycobacte
40	274.2	18.4	4795	2	AAQ51531	M. tuberc
41	274.2	18.4	4795	2	AAI16969	M. tuberc
42	271.8	18.3	2154	8	ACA21214	Prokaryot
43	271.8	18.3	2181	9	ADA32702	DNA encod
44	259.4	17.4	2391	8	ACA23469	Prokaryot
45	253.6	17.0	2202	13	ADT41575	Bacterial

ALIGNMENTS

RESULT 1

AAZ36101

ID AAZ36101 standard; DNA; 1489 BP.

XX AAZ36101;

DT 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

XX IS91; ds.

OS Escherichia coli.

XX WO955908-A2.

PD 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

XX Escherichia coli, particularly serotype O157:H7, used for detecting these

XX Bacteria in food.

XX Claim 1; Fig 1; 48pp; French.

The present sequence is specific to enterohemorrhagic Escherichia coli (EHEC). The sequence is 99.9% homologous to the katP gene of E. coli O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of the present sequence). The present sequence is of plasmid origin. Fragments of the present sequence are used, as probes and primers, for detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or animal samples, foods or the environment. The fragments are also useful

CC for epidemiological studies

[illegible]

isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 42350; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 2214 BP; 662 A; 461 C; 557 G; 534 T; 0 U; 0 Other;

Query Match 35.0%; Score 521; DB 8; Length 2214;

Best Local Similarity 71.7%; Pred. No. 2.le-155;

Matches 683; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

```
QY 537 ATGATAAAAAAACTCTTCGTGTTCTGATCTTCGGCGCTATCGGGAGCTTTCTACC 596
DB 1 ATGTTAAAAAAATCTTACCGTGTAATTAATCTTCGCCATGTACATAATACACCTACG 60
QY 597 GCTGTAGCCGTGATAAAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTCGATTTA 656
DB 61 GCTTGGCAGCAGAACGCCAACACTGATAGTTTCTTACTTACCCAAAGCCTCGATCTC 120
QY 657 ACTCCTCTGAGATTACACAGCCCTGGAATCAATCCCTGGGGGGCTGATTTTGTATATGCC 716
DB 121 TCTCCACTCCGTTTACATAATATAGAGTCCAAATCCTTACGGAAGAGATTTTAATTATGCT 180
QY 717 ACCAGATTTCACAGCTGGATATGAGGCTCTGAABAAAGATATCAAGATTTGCTGACA 776
DB 181 CAACAGTTTAAACACCTGGATCTTGAAGCAGTAAGAAGAGATATAAAACAGTCTTACC 240
QY 777 ACTTCCAGGATTTGGTCCCTCGCGATTATGGTCATTATGGTCCTTTCTTTATTCGTATG 836
DB 241 ACATCACAAGACTGGTGGCCCTGCTGATATGTTATGTTATGTTCCATTTCTTTATTCGTATG 300
QY 837 GCTTGGCAGCGTCCGGAACATACAGGACATATGATGGCCGGGAGCGCCAGTGGTGGT 896
DB 301 GCGTGGCATGTTGGCGGAACCTTACCGCATATATGATGGCCGTGGTGGTGGCGGCGG 360
QY 897 CAGCAAGCTTTGNAACGCTGACAGCTGGCCGATTAACGTTATCTGGATAAGCCCGT 956
DB 361 CAGCAAGATTGTTGAGCCACTCAATAGCTGGCCAGATAACGCCAACCTTGTATAAGCCCGT 420
QY 957 CGATTGCTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCTCGGGAGACCTGATG 1016
DB 421 CGGCTTCTGTGGCCTATATAAAGAAATATGGCCCTAAATATCTGTTGGGCGACCTGATG 480
```

```
QY 1017 GTCTGACTGTAATGTTGCCCTTGAATCCATGGGATTTAAACCGCTGGATTTGCTGGC 1076
DB 481 GTCTTACAGGCAATGTCGCGCTCGAATCTATGGGTTTAAACCGCTGGGTTTTCAGGA 540
QY 1077 GGAAGAGAAGATCACTGGGAGTGGGACTCGGTATCTACTGGGGGCTGACAAACAGCTCTT 1136
DB 541 GCGCGTGAGGATGATTTGGCAATCGATCTGCTTACTTGGGGGGGAGGCAACAGATGCTG 600
QY 1137 GCAGATAACCGGATAAAAACGGGAACTTTAGAAACCTTTGCGCCACGACGATGGGA 1196
DB 601 TCTGATAACAGGGATAAAAATGCAAGCTACCCAAACCGCTGGCGGCGACACAAATGGG 660
QY 1197 CTATTATATCTCAATCTGAAGCCCGGTGGAACCAAGATCTCTGGCTTCCGCGAAA 1256
DB 661 TTGATTTTACCTAAACCTTGAAGGACCAATGGTAAACCAAGCCGTTGCTGGCGCAAAA 720
QY 1257 GATATCAGGGAAGCTTTTTCACGTATGGCCATGATGAGGAGACTGTGGCCCTGATC 1316
DB 721 GATATTGCTGAGGCTTTTCCAGGATGGCAATGATGAGGAGACCGTGGCATTGATT 780
QY 1317 GCGGAGGGCATACATTTGTTAAAGCAATGTTGTCAGCTCTCTGAAAAATGTTATGGC 1376
DB 781 GCGGAGGGCATACCTTTCCGTTAAAGCGCATGGCGCAGCTCTCTGAAAAATGCTGGGT 840
QY 1377 GCAGGCGCTGATGTCACCTGTGGAGGACGAGGACTGGATGGGAAAAATTAATGTGGT 1436
DB 841 GCAGCACCCGGTGAAGCGGGCTAGAACAAACAGGTTTAGGATGGGCGCAATAAATGTGGT 900
QY 1437 ACAGGAAACGGCAAAATATACCATCACCAGTGGCTTGAAGGAGGCTGTGTCGAC 1489
DB 901 TCAGGTATGTTAGGACACGATAACCGTGGGTTGGNAGGGGCGATGACAAAC 953
```

RESULT 3

ADT44922

ID ADT44922 standard; cDNA; 2205 BP.

AC ADT44922;

XX 02-DEC-2004 (first entry)

DE Bacterial polynucleotide #19673.

XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
PS Claim 1; SEQ ID NO 43360; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2205 BP; 500 A; 570 C; 675 G; 460 T; 0 U; 0 Other;

Query Match 26.5%; Score 394.6; DB 13; Length 2205;
Best Local Similarity 66.6%; Pred. No. 7.1e-115;
Matches 581; Conservative 0; Mismatches 289; Indels 3; Gaps 1;

QY 620 GACTCAAAATTTCTACTATCCAGAAACACTGGATTAACTCCTGAGATTACACAGCCC 679
DB 81 GATGACAGTTTTTGGTGGCGGATCACTGGATCTGAGACCGTTGGCCAGAAATTCGGC 140
QY 680 TGAATCAAAATCCCTGGGGGGCTGATTTTGATTTATGCCACCAGATTTCAACAGCTGGATAT 739
DB 141 TGAGTCTAACCCGTTGGCGAAAGCGTTCAACTATGCGGAACAGTTTAAACCGCTCAACCT 200
QY 740 GGAGGCTCTGAAAACATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGC 799
DB 201 CAAAGCGGTGAAAGAGATATCGGAACGTCGTCATCTCCAGCCTTTGGTGGCGGC 260
QY 800 GGATTTGGTCAATATGCTCTTCTTTATTTGATGCTTGGCAGCGTTCGCGGAACATA 859
DB 261 GGACTACGCAACTACGGACCATTTGTCATTCGATGCGCTGGCATAGTGGCGCGGTGA 320
QY 860 CAGGACATATGATGGCGGGAGCGCCAGTGGTGGTTCAGCAACGTTTGAACCGCTGAA 919
DB 321 CCGCATTTTCGACGAGCTGTGGCGCTTCGCGGGATAGCAACGCTTCGAACCGCTCAA 380
QY 920 CAGCTGCCGCGATCACTTAATCTGGATAAGCCGTCGATGCTGTGGCCAGTCAAGAA 979
DB 381 CAGTGGCCGGATAAATCGTCAATTCGGAAGCGCGCGCTTGTATGGCGGATCAATA 440
QY 980 AAAATACGGTCCAGTATTTCTCGGGGAGACCTGATGCTCTGACTGCTAATGTGGCCT 1039
DB 441 GAAATACGGCAGCAGCTTTCTGGGCTGATCTGATGCTGTGGCCGCAATGTGGCGTT 500
QY 1040 TGAATCCATGGATTTAAACCGCTGGGATTTGCTGGCGGAAGAGAAGATGACTGGGATC 1099
DB 501 GGAGGATATGGGATTTCAAAACCATAGGTTTTCGGGTGGCGCAGCGGATGTTGGGAAGC 560
QY 1100 GGACCTGGTATATCGGGGGCTGACACAGCCCTTTCGAGATACCGG---GATAAAAA 1156
DB 561 CGAGATCGTCAACTGGGGAAACAGAAAAGGAATTTCTCGCCGATGAGCGCCATGATAAAG 620
QY 1157 CGGGAACCTTTCAGAAACCTCTTGGCCGCGCAGCAGATGGGATTTATTTATGCAATCTGA 1216

DB 621 AGTGAACCTGGCGAAGCCTCTGGCTCCGTCAGATGGGCTGATCTACCTCAACCCGGA 680
QY 1217 AGCCCCCGGTGGAAAACACAGATCCTCTGGCTTCGCGGAAAGATATCAGGGAAGCTTTTC 1276
DB 681 AGGCGCGGCGGTAAACCCGATCCGTTGGCAGCTGCCAGCATATCCGCAATCCTTCGG 740
QY 1277 ACCTATGGCCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATACATTTGG 1336
DB 741 CCCTATGGCGATGAATGACGAAGAGCGGTGGCTCTGATCGCGGTGGCCATACCTTCGG 800
QY 1337 TAAAGCACATGTGTGCAGCGTCTCTGAAAATATTTGGCGCAGGGCCTGATGGTGACC 1396
DB 801 CAAGGCGCACGGGCGCATAAAGCCGGAAGATCGTTGGCAAGAAACCGCGCGCTGG 860
QY 1397 TGTGGAGGACGAGGACTGGATGGAAAATAATGTTGTTACAGGAACCGCAATATAC 1456
DB 861 TATCGAGGAGCAAGGTTTGGTTTGGGCCAATAAATGTTGGCAGCGGCTATGGCGTCGATAC 920
QY 1457 CATCACAGTGGCCTCGAAGGAGCCTGGTGCAC 1489
DB 921 AGTCACAGTGGTCTGGAAGGGCGTGGTCAC 953

RESULT 4

ACA37119

ID ACA37119 standard; DNA; 1358 BP.

XX ACA37119;

AC ACA37119;

XX 19-JUN-2003 (first entry)

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #18776.

DE ACA37119 standard; DNA; 1358 BP.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

KW Legionella pneumophila.

XX WO200277183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR P-PSDB; ABU33249.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 24989; 1766pp; English.

PS The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1358 BP; 381 A; 313 C; 346 G; 318 T; 0 U; 0 Other;

Query Match 25.7%; Score 392; DB 8; Length 1358;
Best Local Similarity 66.8%; Pred. No. 5.9e-111;
Matches 576; Conservative 0; Mismatches 280; Indels 6; Gaps 2;

628 ATTCTACTATCCAGAACACTGGATTAACTCTCTGAGATTACACAGCCCTGAATCAA 687
D 5 AATACTGGTGCCTTAAATGCTCGATCTTACCCGCTCGGCAACCCCAATGCCACTTGA 64
Q 688 ATCCCTGGGGGGCTGATTGATTATGCCACCAGATTTCACAGCTGGATATGAGGGCTC 747
D 65 ATCCCATGGGTGAAAAATTCACATGTCGGAAGAATTCATAGCTTAGATTGAATGCG 124
Q 748 TGAATAAGATATCAAGATTGCTGACAACTTCCAGGATTGTGCTCCGGAATATG 807
D 125 TGATAGAAGATCTCAAAAAATTAATGACTACTACGCAAGACTGGTGGCTGCTGATTATG 184
Q 808 GTCAATTATGGCTCTTTTATTCGTATGGCTTGGCAGCGTCCGGAACATACAGGACAT 867
D 185 GTAATATGTTCCATTAATCATAGCATGTATGGCATGCGGAGAACTTACCGCATCT 244
Q 868 ATGATGGCGGGAGGCGCCAGTGGTGTGACAGCAACGTTTGAACCGCTGAACAGCTGGC 927
D 245 ATGATGGCGGTGGAGCAACCGTGGCTTCCAGCGTTTCGCCCGCAAAACAGCTGGC 304
Q 928 CGGATAAGCTTAATCTGGATGAACCGCTGATTTGCTGTGGCCAGTCAAGAAAAATACG 987
D 305 CGGATAATGCCAATCTGGACAAAGCCAGACGTTTACTGTGGCCAAATTAAGCAGAAATATG 364
Q 988 GCTCCAGTATTTCTGGGGAGACCTGATGGTCTGACTGGTAATGTGGCTTGAATCCA 1047
D 365 GGCCAGAGATTTCAATGGCTGATTTGTTGGTACTTGTCTGGAATATGTCGATGAATCTA 424
Q 1048 TGGGATTTAAACGCTCGGATTTGCTGCGGAAGAGAGATGACTGGGAGTTCGACCTGG 1107
D 425 TGGGCTTCAAGACCAATGGATTGCTGGAGGGCGTGAAGACGATGGAGGCCATCAATA 484
Q 1108 TATCTGGGGCTTGACAAACAGCCCTCTTGAGATTAACCG---GGATAAAACCGGAAAC 1164
D 485 TCAATTGGGGCCCTGAAAGGCAATAGGTTGGAAGTAAAGCGCCAGGATAAAGATGGAAC 544
Q 1165 TTCAAGAACCTCTGCGCCAGCAGATGGGACTTATTATGTCATCTGAGGCGCCG 1224
D 545 TTGAAAACCGCTGCTGCGACCGTAATGGGCTTAATCTATGTAATCCGGAAGGACCA 604
Q 1225 GTGAAAAACAGATCTCTGCTTCCGGAAGATATCAGGGAAGCTTTTTCAGCTATGG 1284
D 605 ACGGCGTTCAGATCTCTGCTGCGGCGGCAAAAAATTCGCGAGACTTTCGGGCTATGG 664

QY 1285 CCATGGATGATGAGGAGACTGTGCGCCCTGATCGCGGAGGCGCATACATTTGGTAAAGCAC 1344
D 665 CCATGAATGACGAAGAAACAGTTGCTTTAATTTGGGGCGGCGACGATTCGGAATAACAC 724
QY 1345 ATGTTGACGCTCTCTTGAATAATGTTATTTGGCGAGGCGCTGATGTCACCTGTGGAGG 1404
D 725 ATGAGAGCGCT---CCGGCAATATCTGGGGCCAGCCAGAGAGCAGCGCATAGAAG 781
QY 1405 AGCAGGACTGGGATGGAATAATAATGTTGTACAGGAAACGGCAAAATATACCATCACCA 1464
D 782 AACAAGGCTTTGGTTGGAATAACAGTTACGCGAGCGGCAAGGGAAGACACGATTACCA 841
QY 1465 GTGCGCTGGAAGAGCGCTGGTC 1486
D 842 GTGGTTTAGAAGGTGCTGGAC 863
RESULT 5
ADS58848
ID ADS58848 standard; cDNA; 2139 BP.
XX
AC ADS58848;
XX
DT 02-DEC-2004 (first entry)
DE Bacterial polynucleotide #10835.
DE Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
FN US2002333675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 34522; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests.
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at [seqdata.uspto at seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 2139 BP; 406 A; 735 C; 685 G; 313 T; 0 U; 0 Other;

Query Match 25.2%; Score 375; DB 13; Length 2139;
Best Local Similarity 64.5%; Pred. No. 1.3e-108;
Matches 577; Conservative 0; Mismatches 315; Indels 3; Gaps 1;

Qy 595 CCGCTGTAGCGGTGATAAAGAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATT 654
Db 5 CCGCAACCGCGCTCCACCCCAATGTCGAACAGGACTGTGGCCCAATGCGCTCGACC 64
Qy 655 TAACCTCTCTGAGATTACACAGCCCTCAATCAAAATCCCTGGGGGCTGATTTGATTATG 714
Db 65 TTTCGCCCTTCGCCACAGCGGTTGAAATCGAACCCGATGGGCGCAAGTTCAACTATG 124
Qy 715 CCACACAGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAAAGATTGCTGA 774
Db 125 CCGAGGAATTCAGACTCTCGACCTCGCCGCGGTGAAGAGGACATCGAGCGCTGATGA 184
Qy 775 CAACCTCCAGGATGTGCGCTCGGATATATGTCATATATGTCCTTTTATTTCGTA 834
Db 185 CGACCTCGAGGACTGTGGCGCGCGCACTACGCCCACTACGCTCCGTTCTTCAATCCGGA 244
Qy 835 TGGCTTCGCGCGGTGCGGGAACATACAGGACATATGATGCGCGGAGCGCCAGTGGTG 894
Db 245 TGGCATGGCACGCGCGGACCTATTCGACCGCGGAGCGGCGCGGTGCGCGCGCG 304
Qy 895 GTACAGCAAGCTTTTGAACCGCTGAAACAGCTGGCGGATAACGTTAATCTGGATAAAGCCC 954
Db 305 GCCACGAGCGCTTCGAACCGCTCAACTCTTGGCCCGACACGTCACCTCGACAAGGCC 364
Qy 955 GTCGATTTCTGTGCGCAGTCAAGAAAAATACGCTCCAGTATTTCTTGGGAGACCTGA 1014
Db 365 GCCGTCTGCTGTGCCCGATCAAGCAGAGTACGCTCGCAAGATCTCGTGGCGGACCTCA 424
Qy 1015 TGGTCTCTGACTGTAATGTTGSCCTTGAATCCATGGGATTTAAAAACGCTGGGATTTGCTG 1074
Db 425 TGGTGTGACTGGCAACGTCGCGCTCGAATCGATGGGCTTCAGACCTTCGGGTTGCGG 484
Qy 1075 GCGAAGAGAGATGACTGGGAGTCCGACCTGGTATATCTGGGGCCTTGACAAAGCCTC 1134
Db 485 GTGCGCGCGCGAGCTGGGAGGCGCATCAGGCTCTTCTGGGGGCCAGAGAACAAAGTGGC 544
Qy 1135 TTGCAGATAACCGGGATAA---AAACGGGAACCTTCAGAAACCTTTGCGCCACGAGAGA 1191
Db 545 TGGCCGACCGCGCTACCAACGCGACCGGAGCTCCAGAACCCGCTCGCAGCGGTGCAAA 604
Qy 1192 TGGGACTTATTATGTCATCTCAAGCGCCCGGTGGAAAAACAGATCTCTCTGCTTCGG 1251
Db 605 TGGGCTCATCTAGTCAATCCGAAGGCCCGAAGCGCAATCCGACCGCTGTCTCGGG 664
Qy 1252 CGAAAGATATCAGGGAAGCTTTTTCAGTATGGCCATGGATGATGAGAGACTGTGCCCC 1311
Db 665 CAAGGACATCCGCGAGAGGTTTCGCGCGCATGTCGACGACGAGAGACCGCTCGCCC 724
Qy 1312 TGATCGCGGAGGGGATACATTTGGTAAAGCACATGGTGAGCGTCTCTCTGAAAAATGTA 1371
Db 725 TGATCGCGGGGGGCCACACTTTCGCGAAGGCGCACGCGCGCGCAAGCCGAGGCTGCG 784
Qy 1372 TTGCGCGAGGCGCTGATGTTGCACTGTGTGAGGAGCAGGGACTGGGATGGAAAAATAAT 1431

Db 785 TGGGCGTCGATCCGGCGCGCGAGCGTTCGAGCACCGAGGCTCTGGGCTGGAACAAGT 844
Qy 1432 GTGGTACGGAACCGCAATATACCATCACAGTGCCTGGAGGAGCCTGGTC 1486
Db 845 GCGGCAAGGGCAATGCCGAAGATACCGTGAAGCAGCGGCTTCGAAGGCGCGTGGAC 899

RESULT 6

AAV06555
ID AAV06555 standard; DNA; 2238 BP.

XX AC AAV06555;

XX DT 03-JUL-1998 (first entry)

XX DE Microscilla furvescens catalase-53CA1 gene.

XX KW Catalase; epoxidation; hydroxylation; biosensor; paper bleaching;
pasteurisation; ss.

XX OS Microscilla furvescens.

XX FH Key Location/Qualifiers

FT CDS 1..2238

FT FT /*tag= a

FT FT /product= "Catalase-53CA1"

XX WO9800526-A1.

XX PD 08-JAN-1998.

XX PF 03-JUL-1997; 97WO-US016513.

XX PR 03-JUL-1996; 96US-00674887.

XX PA (RECO-) RECOMBINANT BIOCATALYSIS INC.

XX PI Robertson DE, Sanyal I, Adhikary RS;

XX DR WPI; 1998-086953/08.
P-PSDB; AAW33810.

XX PT New bacterial catalases, related nucleic acid vectors and transformed
cells - used as oxidising agents and for detecting or destroying hydrogen
peroxide, e.g. in biosensors.

XX PS Claim 3; Fig 2; 35pp; English.

XX The present sequence is of the Microscilla furvescens catalase-53CA1
gene. Fragments of the gene can be used to identify related sequences.
CC Catalase-53CA1 may be used to catalyse oxidation reactions such as
CC epoxidation or hydroxylation. The enzyme can also be used to detect or
CC destroy hydrogen peroxide, e.g. in connection with glycolic acid
CC production, biosensors, contact lens cleaning, pulp/paper bleaching and
CC pasteurisation of dairy products. Antibodies raised against catalase-
CC 53CA1 can be used to screen libraries for detection and purification of
CC cells containing the enzyme

XX SQ Sequence 2238 BP; 634 A; 545 C; 605 G; 454 T; 0 U; 0 Other;

Query Match 24.0%; Score 357.6; DB 2; Length 2238;

Best Local Similarity 66.7%; Pred. No. 5.3e-103;

Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

Qy 683 ATCAAAATCCCTGGGGGCTGATTTGATTATGACCACAGATTTCAACAGCTGGATATGGA 742

Db 168 ATCGACCCAAACGACCCGGATTTTGACTATGCCAAGAGTTTAAAGAGCTAGATCTGGC 227

Qy 743 GGCTCTGAAAAAGATATCAAGATTTGCTGCAACTTCCAGGATTTGGTCCCTGCGGA 802

Db 228 AGCGGTTAAAAAGAGACCTGGCAGCGCTTAATGACAGATTCACAGGACTGGTGGCCACGAGA 287

Db 528 AACTATGGGCTTTAAACATTTTGGTTTTCAGGTGGCAGACAGATGTATGGAGCCTGA 587
Qy 1100 GGACCTGGTATCTGGGGCTGACACAGACCTTTGTCAGATAACCGGATAAA---AA 1156
Db 588 AGAAGATGTATCTGGGGAGCAGAAACCGAATGGCTGGGAGACAAGCGCTATGAAGTGA 647
Qy 1157 CGGGAACCTTCAGAAACCTTTGGCGGCAGCAGATGGGACTTATTTATGTCAATCCTGA 1216
Db 648 CCGAGAGCTGAAATCCCTGGGAGCGGTACAAATGGGACTCATCTATGTAAACCGGA 707
Qy 1217 AGGCCCGGTGGAACACAGATCCTCTGGCTTCGCGAAGATATCAGGAAGCTTTTTC 1276
Db 708 AGGACCCACCGCAAGCAGACCTATCGCTGCTGCGGTGATATTCGTGAGACTTTGG 767
Qy 1277 AGTATGCCATGATGAGGAGATGTGGCCTGATCGCGGAGGGCATACATTTGG 1336
Db 768 CCGAATGGCAATGAATGACGAAGAAACCGTGGCTCTCATAGCGGTGGACACACCTTCGG 827
Qy 1337 TAAAGCACATGGTGCAGCGTCTCTCGAAATAATGTTGGGCGCAGGGCTGTATGGTGCAAC 1396
Db 828 AATAACCATGGTCTGCCGATGGGAGAAATATGTGGGCGGAGAGCTGCGCGCGAGG 887
Qy 1397 TGTGGAGGAGCAGGAGCTGGGATGGAATAAATAAATGTTGTACAGGAAACGGCAATATAC 1456
Db 888 TATTGAAGAAATGAGCTGGGGTGGAAAAACACCTACGCGACCGACACGTCGCGATAC 947
Qy 1457 CATCACCAGTGGCTGGAAGAGCCTGGTGA 1488
Db 948 CATCACCAGTGGACTAGAGCGCCTGGACCA 979

RESULT 8

ADSS57655
ID ADS57655 standard; cDNA; 2145 BP.
XX AC
XX ADSS57655;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #9642.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX

PS The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or by
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2145 BP; 484 A; 609 C; 603 G; 449 T; 0 U; 0 Other;

Query Match 23.2%; Score 344.8; DB 13; Length 2145;
Best Local Similarity 66.9%; Pred. No. 6.5e-99;
Matches 538; Conservative 0; Mismatches 257; Indels 9; Gaps 3;
Qy 687 AATCCCTGGGGCTGATTTTGGATTATGCCACAGATTTCAACAGCTGGATATCGAGCT 746
Db 64 AATCCCATGGATCCGGACTTCAATTATGCTGAAGAAATTCAGAACTTGACCTGGCAGCC 123
Qy 747 CTGAAAAAAGATATCAAGAATTTGCTGACAACTTCCACAGATTTGGTGCCTTCGGGATTAT 806
Db 124 GTAAAGAAAGATCTTCTGCTTAATGACCGATTCCACAGATTGGTGCCTTCGGGATTAT 183
Qy 807 GGTCAATTATGGTCTTTCTTTTATTTATGATGGCTGGCAGCGTGGCGGAACATACAGACA 866
Db 184 GGCACATACGGGCTCTCTTCATCCGATGGCTTCGTCGACAGTGGCGGAACATACCGTTG 243
Qy 867 TATGATGCGCGGAGCGCGCCAGTGGTGCAGCAACGTTTGAACCGCTGAACAGCTGG 926
Db 244 AACGACGACCGGGCGCGCGGAGACGGAACCCAGCGCTTTGCTCCCTCAACAGCTGG 303
Qy 927 CCGGATAACGTTAATCTGGATAAAGCCCGTTCGATTGCTGTGGCCAGTCAAGAAAAATAC 986
Db 304 CCGGACATATGTAATCTGGATAGGCCCGCGCTTTGCTCTGGCTATTAAACAGAAATAT 363
Qy 987 GGCTCCAGTATTTCTGGGGAGACCTGATGGTCTCTGATGTTATGTTGCCCTTGAATCC 1046
Db 364 GGAATAAAAAATCTCTGGGCGGATCTCATGTCCTGGCGCAATTTGTGCTTTGGAATCC 423
Qy 1047 ATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAAGATGACTCGGAGTC---GGAC 1103
Db 424 ATGGGCTTCAAGACCTTTTGGCTTCGCGCGCGCGCGAGGATGTTTGGGAACCTCAGNA 483
Qy 1104 CTGGTATACCTGGGGCTCTGACAAACAGCCTCTTTCAGATTAACCGGGATATAAAGCG---GG 1160
Db 484 GATATTTATTTGGGCTCTGAAGGAGATGGCTGGCGCACAGCGCTATTCGGGGATCGG 543
Qy 1161 AAACCTCAGAAACCTCTTTCGGCCACCGCAGATGGGACTTATTTATGTCAATCTCGAAGGC 1220
Db 544 GATCTTGAGAAACCTCTCGCGCGCAGTACAGATGGGCGCTGATTTATGTATTAACCCGGAAGGT 603
Qy 1221 CCCGGTGGAAAAACAGATCTCTGGCTTCCGCGAAGATATACAGGAGCTTTTTCAGT 1280
Db 604 CCCAATGGACAGACCCAGTCTGCTTGGCTTCGCGCGCGCGAGCTGCGGGATACCTTTAAACGT 663

Db 850 GCTCCGATTGAAGCACAAAGCTTAGCTGGCAAAACACATACGCTCTGGCAAGGTGCG 909
Qy 1452 TATACCATCACCATGTCGCTGGGAAGAGCTGTGTC 1486
Db 910 GACACGATCACTAGTGGCTTTGAAGGGCTTTGGAC 944

RESULT 10

ADSS6510
ID ADS6510 standard; cDNA; 2163 BP.

XX AC ADS6510;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #8497.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.

XX OS Bacteria.

XX PN US2003233675-A1.

XX FN 18-DEC-2003.

XX PD 20-FEB-2003; 2003US-00369493.

XX PF 21-FEB-2002; 2002US-0360039P.

XX PR (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 32184; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polynucleotide used in
XX the scope of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2163 BP; 407 A; 648 C; 749 G; 359 T; 0 U; 0 Other;

Query Match 22.2%; Score 330.6; DB 13; Length 2163;
Best Local Similarity 64.1%; Pred. No. 2.4e-94;
Matches 549; Conservative 0; Mismatches 299; Indels 9; Gaps 3;

Qy 639 CCAGAAACACTGGATTAACTCTCTGAGATTACAGACCCCTGATCAAAATCCCTGGGG 698
Db 58 CCCGAGCGCTGAATCTGCATATCTCTGCCGACAGCCCTCGTTGTCCGATCCGATGGGA 117
Qy 699 GCTGATTTTGTATGATGCCACAGATTCAACAGCTGATATGAGGGCTCTGAAAAAAGAT 758
Db 118 GAGCTTTTCGATTATATGGAAGGCGTTTAAAGTCTGACCTCGCGGGGTCAAGAAGAT 177
Qy 759 ATCAAGATTGTGCAAACTTCCAGGATTGTGTCCTCGGATTAATGTCATATATGTT 818
Db 178 CTGNAAGCGCTGATGACCGATTTCGAGTCTCTGTGGCGCGGATTTTCGGGCACCTACGGC 237
Qy 819 CTTTCTTTTATTCGTATGGCTTGGCAGGTGCCGGAACATACAGACATATGATGGCGG 878
Db 238 CCGTTGTTCGTCCGATGGCTTGGCACGCGCAGGTACCTACCGCATCGGCGATGGGCGT 297
Qy 879 GGAGGCGCCAGTGGTGTGAGCAACGTTTGAACCGCTGAACAGCTGCGCCGATAAAGTT 938
Db 298 GCGGTGCGCGCTGCGCAGCAGCGTTTCGCGCCACCAACAGCTGCGCGGACACAGTC 357
Qy 939 AATCTGGATAAAGCCCGTTCGATTGCTGTGCCAGTCAAGAAAAAATACGGCTCCAGTATT 998
Db 358 AGTCTGACAAAGCAGCAGCGCTCATCTGCGCGATCAAGCAGAAATACGGCCGCAAGATC 417
Qy 999 TCCTGGGAGACCTGATGTTCTGACTGTAATTTGCTTCCCTTGAATCCATGGGATTTAAA 1058
Db 418 TCGTGGCGCCAGCTGATCGTTCTGACGGGCAATTTGCTTCCCTGAGTGCATGGGTTCAAG 477
Qy 1059 ACGCTGGGATTTGCTGCGGAAGAGAGATGACTGGGAGTCCGACCTG---GTATACTGG 1115
Db 478 ACCTTCGATTTCGGCGAGAGACGCGAGAGTGTCTATGAGCGGACGAGTCCGTCTACTGG 537
Qy 1116 GGGCTTGACAAACAGCCCTCTTTCAGATAACCGGATA---AAAAACGGGAACTTCAGAAA 1172
Db 538 GGCAATGAAGCCGAGTGGCTGGCGGACAAAGCGTTACAGCGGTAAACCGGAACCTCGAGA 597
Qy 1173 CTTCTGCGCCAGCAGATGGGACTTATTTATGTCATCTGAAGCCCGGTTGGA 1232
Db 598 CCGTGGCTGCGGTGCAGATGGGCTGATCTATGTAATCCGGAAGGCCCAATGGCAAC 657
Qy 1233 CCAGATCTCTGGCTTCGCGAAGATATCAGGGAAGCTTTTTCACGTATGCGCATGGAT 1292
Db 658 CCGGACCGGTTGCGCGCGCATCGACATCGCGAGAGTTCGCGCGCATGGCCATGAAC 717
Qy 1293 GATGAGGAGACTGTGGCCCTGATCCGCGGAGGGCATACATTTGTTAAAGCATGTGTGA 1352
Db 718 GACGAAGAAACCGTCGCGTGTATCGCGGCGGTATGCTCTTCGCGAAGACGCATGGCGCC 777
Qy 1353 CGGTCTCTGAAAAATGTTATTTGGCGCAGGCTGATGGTGCACTGTGTGGAGGAGCAGGA 1412
Db 778 G---GCCCGCATCGCACGTCGTGGGCGCCGAGCCTGAAGCCGCGGCTCGAGAGAGCGGC 834
Qy 1413 CTGGGATGAAAAATAATGTGTACAGGAACCGGAAATATACCATCACCAGTGGCCCTG 1472
Db 835 CTTGGCTGGCCAGCAGCTTTGGCACCGGCAAGGGCGGTGATGCCATCGGCAGTGGCCCTG 894
Qy 1473 GAAGGAGCTTGGTCGAC 1489
Db 895 GAGTCTCTGACCAC 911

RESULT 11

ADSS60256

ID ADS60256 standard; cDNA; 2271 BP.

XX AC ADS60256;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #12243.

XX KW Recombinant DNA construct; transformed plant; improved plant property;

XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX KW pathogen tolerance; pest tolerance; plant disease resistance;

XX KW cell cycle pathway modification; plant growth regulator;

XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;

XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX KW bacterial polynucleotide; gene; ss.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPT; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide

XX PT for expression of a polynucleotide encoding a polypeptide from a

XX PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 35930; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a

XX CC promoter functional in a plant cell, where the promoter is positioned to

XX CC provide for expression of a polynucleotide encoding a polypeptide from a

XX CC microbial source. The invention also relates to a transformed plant

XX CC comprising the recombinant DNA construct and a method of producing a

XX CC transformed plant having an improved property. The plant is a crop plant

XX CC such as maize or soybean. The method of producing a transformed plant

XX CC having an improved property comprises transforming a plant with the

XX CC recombinant DNA construct and growing the transformed plant, where the

XX CC polynucleotide or polypeptide is useful for improving plant properties.

XX CC The recombinant DNA construct is useful for producing plants with

XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,

XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

XX CC increased resistance to plant disease, better growth rate by modification

XX CC of the cell cycle pathway with plant growth regulators, increased rate of

XX CC homologous recombination, modified seed oil or protein yield and/or

XX CC content, improved yield by modification of carbohydrate, nitrogen or

XX CC phosphorus use and/or uptake, by modification of photosynthesis or by

XX CC providing improved plant growth and development under at least one stress

XX CC condition. Improved lignin production or improved galactomannan

XX CC production. This sequence represents a bacterial polynucleotide used in

XX CC the scope of the invention. Note: The sequence data for this patent did

XX CC not form part of the printed specification but was obtained in electronic

XX CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2271 BP; 439 A; 726 C; 736 G; 370 T; 0 U; 0 Other;

XX Query Match 21.7%; Score 323.2; DB 13; Length 2271;

XX Best Local Similarity 64.6%; Pred. No. 5.7e-92;

XX Matches 532; Conservative 0; Mismatches 283; Indels 9; Gaps 3;

XX 672 CACAGCCCTGATCAAAATCCTGGGGGGCTGATTTTATGTCACACAGATTTCACAG 731

Db 190 CAGTCCTCTCTGTCGACCCGATGGGTGAGCGTTTCGATATGCCGAGGAATTTCAAGAGC 249

Qy 732 CTGGATATGGAGGCTCTGAAAAAAGATATCAAGATTTCGTCACAACTTCCAGGATTGG 791

Db 250 CTCGACCTCGACGCGCTCATCAAGACCTGCACGCGTGATGACGGATTTCGACGAGTGG 309

Qy 792 TGCCTCGGGATTATGTCTATTATGGTCTTTCTTTTATTCGTATGGCTTGGCAGGTGCC 851

Db 310 TGGCGGCGGATTTTCGGCCACTACGGCCGCTGTTCATTCGGATGGCTGCGACACGCA 369

Qy 852 GGAACATACAGGACATATGATGGCGGAGGCGCCAGTGTGTGACAGACGTTTTCGAA 911

Db 370 GGCACCTACCGCATTCGCGACGCGCGCGCGCGCGCGCTGCGCAGCAGCGTTTCGG 429

Qy 912 CCGCTGAACAGCTGGCCGATTAACGTTAATCTGGATAAAGCCGTCGATTCTGTGGCCA 971

Db 430 CCGCTCAACAGCTGGCCGCAATGTCAACCTCTCAAGGCGCCCGCGCTTTTGTGGCG 489

Qy 972 GTCAGAAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGGTCTGCTGCTGAT 1031

Db 490 ATCAAGCAGAAATATGGCCGCAAGATCTCTGGGCCGACCTTCTGATCTCTCACCGCAAC 549

Qy 1032 GTTGCCTTGAATCCATGGGATTTAAAAAGCTGGGATTTGCTGGCGAAGAGAGATGAC 1091

Db 550 GTCCGCTGGATTCGATGGGCTTCAAGACCTTTGGCTTTGCCGGCGCGCGCCGACGTC 609

Qy 1092 TGGAGTCCGACCTG---GTATCTGGGGGCTGACAAACAGCCTCTTTCAGATAACCGG 1148

Db 610 TGGAGCTTGACGAGGAGCTTTACTGGGGTCCCGAAGGCAAGTGGCTGGCCGACGAGCGC 669

Qy 1149 GATA---AAAAAGGAAACCTTCAGAAACCTTTCGCCGCGCAGCAGATGGGACTTTAT 1205

Db 670 TACAGCGCGACCGTGCACACCCGCTCGCGCCGCTGCAGATGGGCTGATCTTAC 729

Qy 1206 GTCAATCTCAAGSCCGGTGGAAACACAGATCTCTGGCTTCGCGAAGAGATATCAGG 1265

Db 730 GTCAATCCGAGGACCGAAGCGCAATCCGATCCGCTGGCCGCGCGCGGATATCAGG 789

Qy 1266 GAAGCTTTTTCAGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGG 1325

Db 790 GACACATTTCGCGCTATGGCGATGAACGACGAGGAAACCGTAGCACTCATCGCGCGGC 849

Qy 1326 CATACATTTGGTAAGACATGGTGACGCTCTCTGAAAAATGTTATTTGGCGCAGGGCCT 1385

Db 850 CATACGTTCCGCAAGACCCATGCT---GCGGGTGACGCGAGCCTGTGGTGTGCGAGCG 906

Qy 1386 GATGGTGCACCTGTGGAGGACGAGGACTGCGATGGAAAAAATAAATCTGTTACAGGAAC 1445

Db 907 GAAGGCGCTGATTCGAGCAGCAGGCGCTTTGGCTGGGCGAGCAAAATTCGGCACC 966

Qy 1446 GGCATAATATACCATCACAGTGGCTGGAAGGAGCCTGGTTCGAC 1489

Db 967 GCGGTCACGCCATCGCAGCGCTGCTGAAGTTCATTGGAGCAC 1010

RESULT 12

ADH12940

ID ADH12940 standard; DNA; 2223 BP.

XX AC ADH12940;

XX DT 25-MAR-2004 (first entry)

XX DE Francisella tularensis immunogenic protein 21 DNA, SEQ ID NO:121.

XX KW Immunogenic protein; protective immune response; vaccine;

XX KW genetic vaccine; antibacterial; gene; ds.

XX OS Francisella tularensis.

XX PN WO2004003009-A2.

PD 08-JAN-2004.
XX
PF 26-JUN-2003; 2003WO-GB002718.
XX
PR 28-JUN-2002; 2002GB-00014942.
XX
PA (MINA) UK SEC FOR DEFENCE.
XX
PI Titball RW, Mayers CN, Duffield ML, Miller J, Rowe SC;
XX WPI; 2004-083016/08.
XX p-PSDB; ADH12840.
XX
XX New protein, useful as a vaccine for producing a protective immune
PT response in a mammal against infection by Francisella tularensis, or
PT preventing or treating Francisella tularensis infection in a mammal.
XX
XX Claim 9; SEQ ID NO 121; 217pp; English.
XX
XX The invention relates to 100 potentially immunogenic proteins from
CC Francisella tularensis (ADH12820-ADH12919) and nucleic acids encoding
CC them (ADH12920-ADH13019) which may be used in vaccines. The nucleic acids
CC may be contained within a vaccine, bacterial or plasmid vector. The
CC invention also relates to a pharmaceutical composition comprising the
CC protein or the nucleic acid in combination with a pharmaceutical carrier
CC or excipient. The proteins and nucleic acids are useful in vaccine
CC compositions for producing a protective immune response against
CC Francisella tularensis infection in a mammal, or for preventing or
CC treating Francisella tularensis infection in a mammal. The present
CC sequence represents DNA encoding a Francisella tularensis protein of the
CC invention.
XX
XX Sequence 2223 BP; 756 A; 385 C; 464 G; 618 T; 0 U; 0 Other;
SQ
Query Match 21.6%; Score 321.4; DB 12; Length 2223;
Best Local Similarity 59.3%; Pred. No. 2.1e-91;
Matches 569; Conservative 0; Mismatches 381; Indels 9; Gaps 1;
QY 537 ATGATAAAAAAAGCTTCTGTTCTGATTCCTTCTGGCGGTATCGGGAGCTTTCTACC 596
Db 1 ATGCTAAGAAAAATGTTAACTGCTTTAGGAATGCTGGAATGCTACTAGCTTCTAGCAAT 60
QY 597 GCTGTAGCCCTGATATAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTA 656
Db 61 GCTATCGCAGAAGATACCAACGAAAAAGATATCTTTCACACACAGAGGTAGATTYA 120
QY 657 ACTCCTCTGAGATTACACAGCCCTGAAATCAATCCCTGGGGGGCTGATTTTGATATGCC 716
Db 121 TCACCATGGCAATTTAAATAAGCTTGATAGCCCAATGGATAAAGATTATACTATCAT 180
QY 717 ACCAGATTCAACAGCTGGATATGGAGCTCTGAAAAAAGATATCAAGATTTCCTGACA 776
Db 181 CAAGCTTTCAAAAAAAGTATGATGAAACAGCTTAAAAAAGATATGCAAGATCTTTTAAAC 240
QY 777 ACTTCCAGGATTTGGTCCCTGCGGATTTGGTATGATGCTTCTTTCTTTATTTCGATG 836
Db 241 CAGTCACAAGACTGGTGGCTGCTGCTTTTGGCAATTTATGGTCTCTTTCTTTATTAGACTA 300
QY 837 GCTTGGCAGGTGCGGAAACATACAGGACATATATGCGCGGGAGCGCCAGTGGTGGT 896
Db 301 TCGTGGCATGATGCTGGTACATACAGAAATATATGTCGAGAGGAGCGCTAAATCGTGA 360
QY 897 CAGCAACGTTTGTAAACCGCTGAAACAGCTGGCGGATACGTTAATCTGGATAAAGCCGCT 956
Db 361 CAACAAAGGTTCTCCCTTTAAATAGCTGGCCAGATAATGTTAATCTTTGACAAAGCAAGG 420
QY 957 CGATTGCTGTGGCCAGTCAAGAAAAATACCGCTCCAGTATTTCTCTGGGAGACTGATG 1016
Db 421 CAACCTTTTATGGCCAAATCAACAAAAATATGCTGATGCTGTTTATGCTGATTTGATT 480
QY 1017 GTCCTGACTGGTAATGTTGCCCTTTGAATFCCAATGGGATTTAAACGCTGGGATTTGCTGGC 1076
Db 481 GTTTTAGTGGTACTGTTTCTTTTAGAATCAATGGAAATGAAGCCTATAGGGGTTTGTCTTT 540

QY 1077 GGAAGAGAAAGATGACTGGGAGTTCGGACCTGGTATATCTGGGGGCTTGACAAACAGCCTCTT 1136
Db 541 GGTAGAGAAAGACGACTGGCAAGGTGATACAAACTGGGGACTATCACCTGAAGAGATA 600
QY 1137 GCAGATAACCGGGATATAAAGCGGAAACTTTAGAAAACCTTTGCGCGCAGCAGATGGGA 1196
Db 601 ATGCTCTAGTAATGTAAGAGATGGCAAACTTGTCTCTGCTATACGCCGCAACACAAATGGGG 660
QY 1197 CTTATTTATGTCAATCTGNAAGSCCGGTGGAAAAACAGATCCTCTGGCTTCGCGCAAA 1256
Db 661 CTAATATATGTAATCCAGAAGGTCTGTATGGTAAACCTGATATCAAAAGTGCAGCTAGT 720
QY 1257 GATATCAGGAAGCTTTTTCACGTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATC 1316
Db 721 GAAATTCGTGAGCCCTTCGAGCTATGGGGATGACAGATAAAGAAACTGTGCGCCCTAATT 780
QY 1317 GCGGGAGGCGATACATTTGGTAAAGCACATGTTGGTGCAGCTCTCTCTG-----AAAAA 1367
Db 781 GCAGGCGGTATACATTTGGTAAAACTCATGTGTGCAGTTCCAGAGGATAAAGTCAACAA 840
QY 1368 TGTATTTGGCGCAGGCGCTGATGTCACCTGTGGAGSAGCAGGACTGGGATGGAATAAT 1427
Db 841 GCAATTTGGACCTGCTCTCTGATAAGGGCGCTTATTTGAGCAGCAGGTCTAGGCTGGCAAT 900
QY 1428 AAATGTGTACAGGAAACGGCAATATATACCATCACCAGTGGCCTGGAAGGAGCTGGTC 1486
Db 901 AGTTATGGCACTGGAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 959
RESULT 13
ADT45831
ID ADT45831 standard; cDNA; 2205 BP.
XX
AC ADT45831;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #20582.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 44269; 122pp; English.
PS


```
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2181 BP; 489 A; 567 C; 649 G; 476 T; 0 U; 0 Other;

Query Match      20.5%; Score 305.4; DB 13; Length 2181;
Best Local Similarity 62.1%; Pred. No. 2.8e-86;
Matches 499; Conservative 0; Mismatches 301; Indels 3; Gaps 1;

Qy 684 TCAATCCCTGGGGCTGATTTTGGATATGTCACACAGATTTCACACCTGGATATGGAG 743
Db 157 TCTAACCCACTGGGTGAGGACTTTGACTACCGCAAGAAATTCAGCAAAATAGATTACTAC 216

Qy 744 GCTCTGAATAAGATATCAAGATTTGCTGCAACTTCCAGGATTTGGTCCCTGCGGAT 803
Db 217 GGCCTGAATAAGATCTGAAGCCCTGTTGACAGAACTCTCAACGTTGGTGGCAGCCGAC 276

Qy 804 TATGTCATATTATGTCCTTTCTTTATTCGTATGCTTGGCTGCGACGCTGCGGCAACATACAGG 863
Db 277 TGGGCGAGTTACGCGGCTCTGTTTATTCGTATGCTGCGACGCGCGGGGACTTACCGT 336

Qy 864 ACATATGATGCGGGGAGGCGCCAGTGGTGTGAGCAACGTTTGAACCGCTGACAGC 923
Db 337 TCAATCGATGACGCGGTGGCGCGGCTGCTGCTGAGCAACGTTTGGACCGCTGAACCTCC 396

Qy 924 TGGCCGGATAACGTTAATCTGATAAAGCCGCTGATTGCTGTGCCAGTCAAGAAATAA 983
Db 397 TGGCCGGATAACGTTAAGCTCGATAAAGCGGCTGCTGTTGTGGCCAAATCAACAGAAA 456

Qy 984 TACGGCTCCAGTATTTCTCGGGAGACCTGATGCTGCTGACTGTTAATGTTGCCCTTGA 1043
Db 457 TATGCTCAGAAAATCTCTCGGGGACCTGTTTATCTCTCGCGGTAAAGTGGCGTAGAA 516

Qy 1044 TCCATGGGATTTAAACGCTGGATTTGCTGGCGGAGAGAAAGATGACTGGGAGTGGAC 1103
Db 517 AACTCCGCTTCCGTACCTTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 576

Qy 1104 CTGCTATGATGGGCGCTGACAAAGCCTCTTGCAGATAACCGGATATAAAGCGGAAA 1163
Db 577 CTGATGTTAACTGGGTGATGAATAAGCCTGGCTGACTACCGTCAATCCGGAAGCGGTG 636

Qy 1164 CTTCAGAAACCTCTTGGCGCCACGATGGGACTTATTTATGTAATCTCTGAAGGCGCC 1223
Db 637 GCGAAGCAGCCCTGGGTGCAACCGAGATGGTCTGATTTACGTTAATCCCGAAGGCCG 696

Qy 1224 GGTGAAACAGATCTCTGCTGGCTTCCGGAAGATATCAGGAGCTTTTTCACGTATG 1283
Db 697 GATCAGCGCGCAACCGCTTCTCGCGCAGCAGCTATCCGCGGACCTTCGGCAACATG 756

Qy 1284 GCCATGATGATGAGGAGACTGTGCCCTGATCCGGGAGGCGCATACATTGTTGAAGCA 1343
Db 757 GGCATGAACGACGAAGAAACCGCTGGCGTGAATGCGGGTGTCTACGCTGGGTAAAC 816

Qy 1344 CATGGTCAGCGCTCTCTCTGAATAATGTATTGGCGCAGGCGCTGATGTGCACTGTGGAG 1403
Db 817 CACGCTGCG--GTCCGACATCAATGTAGTCTCTGATCCAGAGCTGCACCGATTGAA 873

Qy 1404 GAGCAGGACTGGATGTAATAATATGTTGGTACAGAAACCGCAATATACCATCAC 1463
Db 874 GAACAAGGTTTAGTTGGCGAGCACTTACGCGCAGCGCGTGGCGCAGATGCCATTACC 933

Qy 1464 AGTGGCTGGAAGAGCGCTGTC 1486
Db 934 TCTGGTCTGGAAGTAGTCTGGAC 956

RESULT 16
ADS60748
ID ADS60748 standard; cDNA; 2166 BP.
XX
AC ADS60748;
XX
DT 02-DEC-2004 (first entry)
XX
```

DE Bacterial polynucleotide #12735.

KW Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

PI WPI; 2004-061375/06.

DR New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

PT Claim 1; SEQ ID NO 36422; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 2166 BP; 486 A; 630 C; 596 G; 454 T; 0 U; 0 Other;

SQ Query Match 20.4%; Score 303.4; DB 13; Length 2166; Best Local Similarity 62.3%; Pred. No. 1.2e-85; Matches 513; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

Qy 672 CACAGCCCTGAATCAATCCCTGGGGCTGATTTTGAATATGACCAAGATTTCACAG 731

Db 100 CACAACCTCGGTCTCTAATCCCTGGATAAGGGCTTTCGACTATCTGCGCGCTCAATAGC 159

Qy 732 CTGGATATGGGCTCTGAAAAAGATATCAAGATTGCTGACAACTTCCAGGATTGG 791

Db 160 CTGACTACTTCGGACTGAGCGGATCTGAGGCACTCATGACAGACTCCAGGACTGG 219
Qy 792 TGGCTTCGGATTATGATGATATGATGCTCTTTCTTTATTCGATGCTGGTCGACGCTGCC 851
Db 220 TGGCGGCGGACTTGGTCACTATGCGGAGTCTCTTTATCCGATGGCTCGGCACAGTGCT 279
Qy 852 GGAACATACAGGACATATGATGCGCGGAGGCGCCAGTGGTCTGAGCAACGTTTTGAA 911
Db 280 GGAACGATATCGCTCTTTCGCGGTTCGCGGCGGTGGCGGACAGGTGACGACGCTTCGCT 339
Qy 912 CCGCTGAACAGCTGGCGGATAGCTTAATCTGATAAAGCCCGTCGATGCTGTGGCCA 971
Db 340 CCGCTCAACAGCTGGCGGATAGCTGAGCTGAGCAAGGCTGCTGCTCTTATGGCC 399
Qy 972 GTCAAGAAAAATACGCTCAGTATTTCTTGGGGAGACCTGATGCTGACTGGTAAT 1031
Db 400 ATCAAGCAGAAGTACGGAAGCAAGATCTCATGGGCTGACTTGTCTGATTCGCGCGGAAT 459
Qy 1032 GTTGGCTTGAATCCATCGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGAC 1091
Db 460 GTGCGCTTGAATCAATGGGGTTCAAGACCTTTGGGTTTGGCGGTGCGCCGAAGTGATACC 519
Qy 1092 TGGAGTCCGACCTGGTATATCTGCGGCGCTGACAAACAA-----GCCTCTTGCAGATAAC 1145
Db 520 TGGAGCAGACCACTGCTGCTCTTGGGAGCGGAGAGGAATGTTGGGTAATGATGTC 579
Qy 1146 CGGGATAAAAACGGGAACCTTCAGAAACCTTTGCGCGCAGCGAGATGGACTTTATAT 1205
Db 580 CGCTACTTGAACGGGAGAACTCGACAAACCGCTTCGCGGCATCACACATGGGTCTTATTTAC 639
Qy 1206 GTCAATCTGAGGCCCGGTGGAAACCCAGATCTCTGGCTTCGCGGAAAGATATCAGG 1265
Db 640 GTTAATCCAGAAAGACCCAAACAGAACCCCGCTTCGCGGCAAGGATATCCGC 699
Qy 1266 GAAGCTTTTTCACGTATGGCCATGATGAGGAGACTGTGGCCCTGATCGCGGAGGG 1325
Db 700 ATCACTTTGGTCAATGGGCATGATGAGGAGACTGTTCCTGATGCTGGTGA 759
Qy 1326 CATACATTTGGTAAAGCAGATGTTGACGCTCTCTGAAATATTTGGCGCAGGCGCT 1385
Db 760 CACAGCTTGGAAAGACGACGG---CGCGGCGCTGCAACCCATCTCGGCAAGAACCA 816
Qy 1386 GATGGTCACTTGGAGGACGAGGACTGGATGGAAAAATAAATGTTGATACAGGAAC 1445
Db 817 CATGGTGGGGTATTGAGTTAAGGCTTAGCTGGGAGACGGCTTCGAGTCTGGGACC 876
Qy 1446 GGCAATATACCATCACCAGTGGCTTGGAGGAGCCTGGTGA 1488
Db 877 GGGCGACATGCTATCACCAGCGGTCTGGAGGTGATCTGGACCA 919

RESULT 17

ID ADT42268

XX ADT42268 standard; cDNA; 2214 BP.

XX AC ADT42268;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #17019.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; BS.
XX Bacteria.
OS
XX
XX
XX US2003233675-A1.

PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOV/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 40706; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a plant with the
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2214 BP; 387 A; 789 C; 730 G; 308 T; 0 U; 0 Other;

Query Match 20.3%; Score 302.2; DB 13; Length 2214;
Best Local Similarity 62.0%; Pred. No. 3.le-85;
Matches 513; Conservative 0; Mismatches 308; Indels 6; Gaps 2;

Qy 663 CTGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTTGTATTATGCCACACGA 722
Db 136 CTGAACACGACGCGCGCGCTCCCAATCCGATGGCGAGGCGTTTCGACTACGCCGAGCG 195
Qy 723 TTTCACACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAGATTTGCTGACACTTCC 782
Db 196 TTCAAGAGCCTCGACCTGGAGCGCGTCTGTCAGCACCTGCACGCGCTGATGACCGACAGC 255
Qy 783 CAGGATTTGGTCCCTCGGATTTATGTTGTCATTTGTCCTTTTATTTCGTATGGCTTGG 842
Db 256 CAGAGTGTGGTGGCGGTGACTTCGGCCACTATTCGGCGGCTTTCATTCGCTGGCTGG 315
Qy 843 CACGGTCCGGAACATACAGACATATGATGCGCGGGAGCGCCAGTGGTGTGACGAA 902
Db 316 CACGCGCGCGACCTTATCGCATCACCGACGCGCGCGCGCGCGCGCGCGCGCGCGAG 375
Qy 903 CGTTTGAACCGCTGAACAGCTGGCCGATTAACGTTATCTTGGATTAAGCCCGCTGATTG 962
Db 376 CGCTTCGCCCCCTGAAACAGCTGGCGCGGACCAACCAATCTCGAACAGGCCCGCGCTG 435


```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,782
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.141US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-3031
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-418-782-1

Query Match      20.0%; Score 298.4; DB 1; Length 2235;
Best Local Similarity 62.4%; Pred. No. 3.6e-96;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

Qy 688 ATCCCTGGGGGCTGATTTGATATGACCAAGATTTCAACAGCTGGATATGAGGCTC 747
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 748 TGAAGAAAGATATCAAGATTTGTCACACTCCAGGATTTGTCCTGCGGATTTATG 807
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 236 TGACCGGGATCATCGAGAGTGTATGACCACTCGAGCGGTGTCGCGCCGCTACG 295
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 808 GTCAATTATGTCCTTTCTTTATTCGTATGCTTGCGACGTTTCAACAGCTGGATATGAGGCTC 867
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 296 GCGACTACGGCCGCTGTTATCCGGATGGGTGGCAGCTGCGGCACTACCGCATCC 355
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 868 ATGATGCGCGGGAGGCGCCAGTGGTGTAGCAAGCTTTTGAACCGCTGAACAGCTGGC 927
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 356 ACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 928 CGGATACGTTAATCTGGATAAAGCCGCTGATTTGTCGTGCGCCAGTCAAGAAAGATACG 987
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 416 CCGAACAAGCCAGCTTGGACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 988 GCTCAGATTTCTTGGGGAGACCTGATGTCCTGACTGTTAATGTTGCTTGAATCCA 1047
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 476 GCAAGAGCTCTCATGGCGGACCTGATTTGTTTCGCGGCACTGCGCGCTGGAATCGA 535
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1048 TGGATTTAAACGCTGGGATTTGCTGCGGGAAGAGATGATGCGGAGTTCGACCTGG 1107
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 536 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCGCGGTCGACAGTGGGAGCCGATGAGG 595
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1108 TATAGTGGGGGCTTGACAAAGCTCTTTCGAGATACCGGGATA---AAACCGGAAC 1164
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 596 TCTATTGGGCAAGGAAGCCACTTGGCTTCGGGCTGAGCGTTTACAGCGTTAAGCGGATC 655
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1165 TTCAGAACCTCTTTCGCGCCACGAGATGGGACTTATTATGTCAATCTGAAGCGCCCG 1224
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 656 TGGAGAACCCGCTGGCGCGGTGCGATGGGGCTGATCTAGTGAACCGGAGGGCGGA 715
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1225 GTGAAAAACAGATCTCTGCTTCGCGGAAGATATCAGGGAAGCTTTTTTCAAGTATGG 1284
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 716 ACGGCAACCCGAGACCCATGCGCGCGCGGTGCGACATTCGCGGAGACGTTTCGCGGCATGG 775
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1285 CCATGGATGATGAGAGACTGCGCCCTGATGCGGGAGGCATACATTTGTATTAAGCAC 1344
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 776 CCATGAACGAGCTGGAACACAGCGGCGCTGATGTCGCGGGTACACTTTCGTTAAGACC 835
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1345 ATGTGACGCTCTCTGAAAAATGTAATTGGCGCAGGGCCCTGATGGTGCACCTGTGGAGG 1404
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 836 ATGGCGCG---GCCCGCCGATCTGGTTCGGCCCGCCGACCCGAGGCTCTCGCTGGAGC 892
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Qy 1405 AGCAGGAGCTGGGATGGAATAAATGTGTACAGAAACGGCAATATACCATCACCA 1464
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 893 AGATGGGCTTGGCTGGAGAGCTCGTATGGCACCGGACCGGTAGGACCGCATCACCA 952
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1465 GTGGCCTGGAAGGAGCTGTGTCGA 1488
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 953 GCGGCATCGAGGTCGTATGGACGA 976
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-08-228-662-1
; Sequence 1, Application US/08228662
; Patent No. 568639
; GENERAL INFORMATION:
; APPLICANT: COCKERILL, FRANKLIN R.
; APPLICANT: KLINE, BRUCE C.
; APPLICANT: UHL, JAMES R.
; TITLE OF INVENTION: DETECTION OF ISONIAZID RESISTANT STRAINS
; TITLE OF INVENTION: OF M. TUBERCULOSIS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCHWEGMAN, LUNDBERG & WOESSNER, P.A.
; STREET: 3500 IDS CENTER
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/228,662
; FILING DATE: 18-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WOESSNER, WARREN D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.123US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-228-662-1
```

```
Query Match      20.0%; Score 298.4; DB 1; Length 2235;
Best Local Similarity 62.4%; Pred. No. 3.6e-96;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

Qy 688 ATCCCTGGGGGCTGATTTGATATGACCAAGATTTCAACAGCTGGATATGAGGCTC 747
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 748 TGAAGAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGTCCTGCGGATTTATG 807
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 236 TGACCGGGATCATCGAGAGTGTATGACCACTCGAGCGGTGTCGCGCCGCTACG 295
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 808 GTCAATTATGTCCTTTCTTTATTCGTATGCTTGCGACGTTTCAACAGCTGGATATGAGGCTC 867
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 296 GCGACTACGGCCGCTGTTATCCGGATGGGTGGCAGCTGCGGCACTACCGCATCC 355
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 868 ATGATGCGCGGGAGGCGCCAGTGGTGTAGCAAGCTTTTGAACCGCTGAACAGCTGGC 927
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 356 ACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db	654	GGAGAACCCCGCTGCGCCGCGTGCAGATGGGGCTGATCTACGTGAACCCGAGCGCGCGAA	713
Qy	1226	TGGAAAAACAGATCCTCTGGCTTCCGCGAAAGATATCAGGGAAAGCTTTTTCAGGTATGGC	1285
Db	714	CGGCAACCCCGGACCCCATGSCCGCGGCGTGCATTCGCGAGACGTTTCGGCGCATGGC	773
Qy	1286	CATGGATGATGAGGAGACTGTGGCCCTGATCGCGGGAGGGGCATATCTTGGTAAAGCAC	1345
Db	774	CATGAACGAGCTCGAAACACGCGGCGCTGATCGTCGGCGGTACACTTTCGGTAAAGACCCA	833
Qy	1346	TGCTGCAGCGTCTCCTCGAAAAATGTTATTCGCGCAGGGCCTGATGGTCACCTCTGGAGGA	1405
Db	834	TGSGCGCG----GCCCGCGCATCTGCTCGGCCCCGGAACCCGAGGCTGCTCCGCTGGAGCA	890
Qy	1406	GCAGGGACTGGGATGGAAAAATAATGTGGTACAGGAAACCGGCAAAATATACCATCACCCAG	1465
Db	891	GATGGGCTTGGGCTGGGAAGAGCTCGTATGSCCGCGAACCCGGTAAGACGCGATCACCCAG	950
Qy	1466	TGGCCTGGGAAGGAGCCTGCTCGA	1488
Db	951	CGGCATCGAGTCTGTATGGACGA	973

RESULT 15

US-08-852-219-2
; Sequence 2, Application US/08852219
; Patent No. 5922575
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: Of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mueiting, Raasch & Gebhardt, P.A.
; STREET: 119 No. 5922575th Fourth Street, Ste. 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA

Query Match 18.4%; Score 274.2; DB 2; Length 2221;
Best Local Similarity 61.4%; Pred. No. 2.1e-87;
Matches 493; Conservative 0; Mismatches 303; Indels 7;
QY 688 ATCCCTGGGGGGTGAATTTGATTATGCCACCATTTTCAACAGCTGGGATG

Db	176	ACCCGATGGGTGGCGGTTTCGACTATGCGCGGAGAGTCCGACCACTCGACTTCGACCTTC	235
Qy	748	TGAAAAAGATATCAAAAGATTTTCTGACAACTTCCAGGATTGGTGCCTTCGCGATTATG	807
Db	236	TGACGGGGNCATCGAGGAAGTATGACCACTTCGACGGTGGTGGCCGCGCACTACG	295
Qy	808	GTCAATTATGGTCTCTTTCTTATTTCGTATGGCTTGGCACGGTGC CGGAACATACAGGACAT	867
Db	296	GCCACTACGGCGCTGTTTATCCGGATGGGTGGCGACGCTGCCGCGCACTTACCGCATCC	355
Qy	868	ATGATGGCGGGGAGCGGCAGTGTGTGTAGCAACGTTTTGAAACCGCTGAAACAGCTGGC	927
Db	356	ACGACGGCGCGCGCGCGCGCGCGCGCATGCAGCGGTTCCGCGGCTTAAACAGCTGGC	415
Qy	928	CGGATAACGTTAATCTGGATAAAGCCCGTCGATTGCTGTGGCCAGTCAAGAAAAATACG	987
Db	416	CCGACAACCCAGCTTGTGACAAAGGCGCGCGGCTGCTGTGGCGGTCGCAAGAAAGTACG	475
Qy	988	GCCTCCAGTATTTCTCGGGGAGACCTGATGCTGCTGACTGATAA--TGTTGCCCTTCAATC	1045
Db	476	GCAAGAGCTCTCATGGCGGACCTGATTGTTTTTCCCGGCAACCGCTCGGCTCGGAATC	535
Qy	1046	CATGGGATTTAAACCTCGGATTTGCTGCGGGAAGAGAAGATGACTGGGAGTCGGACCT	1105
Db	536	GATGGGCTTCAAGACGTTTGGGTTTGGCTTCGG--GCGTCGACCACTGGGAGACCGATGA	593
Qy	1106	GGTATACTGGGGCTTGCAACAAGCGCTCTTGCGATAACCGGGATAAAAACGGGAACT	1165
Db	594	GGTCTATTGGGCAAGGAAGCCACTGGCTCGCGGATGACGGTTACAGCGTAAGCGATCT	653
Qy	1166	TCAGAACTCTTTGCGCCACGCGAGATGGGACTTATTTATGCAATCTCTGAAGGCCCGG	1225
Db	654	GGAGAACCCGCTGGCCGCGGTGCAGATGGGGCTGATCTACGTGAACCCGAGGCGCGAA	713
Qy	1226	TGGAAAAACGATCTCTCGCTTCGCGGAAAGNATCAGGGAAGCTTTTTTCAGTATGC	1285
Db	714	CGGCAACCCGGAACCCATGGCGCGCGCGTTCGACATTCGCGAGACGTTTTCGGCGCATGGC	773
Qy	1286	CATGGATGATGAGGACATGTGCGCCTGATTCGCGGAGGGCATACATTTGGTAAACACA	1345
Db	774	CATGAACGACGTGGAACAGCGCGCTGATCTGCGCGGTACACTTCGTAAGACCCA	833
Qy	1346	TGTTGCGAGCTCTCTTGAAAAATGTATTGGCGCAGGGCCTGATGTGTGCACCTGTGGAGA	1405
Db	834	TGGCGCG--GCCCGCGATCTGTTGCGGCCCGNACCCGAGGCTGCTCCGCTGGAGCA	890
Qy	1406	GCAGGCACTGGGATGGAAAAATAAATGTGGTACAGGAAACGGCAAAATATACCATCACCA	1465
Db	891	GATGGCTTTGGGCTGGAAGAGCTCGTATGSGCACCGGAACCGGTAAAGGACGCGATCACC	950
Qy	1466	TGGCCTGGGAAGAGCCCTGTGCA	1488
Db	951	CGGCATCGAGGCTCGTATGACGA	973

RESULT 16

```

US-08-459-499-8
? Sequence 8, Application US/08459499
? Patent No. 5871912
? GENERAL INFORMATION:
? APPLICANT: Heym, Beate
? APPLICANT: Cole, Stewart T.
? APPLICANT: Young, Douglas B.
? APPLICANT: Zhang, Ying
? TITLE OF INVENTION: Nucleic Acid
? TITLE OF INVENTION: For Detection
? TITLE OF INVENTION: (Amended)
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSES: Finnegan, Henderson
? ADDRESSES: Dunner
? STREET: 1300 I Street, N.W.
? CITY: Washington

```

```
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,499
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/875,940
; FILING DATE: 30-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/929,206
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/029,655
; FILING DATE: 11-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0110-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4794 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-459-499-8

Query Match      18.48; Score 274.2; DB 2; Length 4794;
Best Local Similarity 61.4%; Pred. No. 3.9e-83;
Matches 493; Conservative 0; Mismatches 307; Indels 7; Gaps 3;

Qy 688 ATCCCTGGGGGGCTGATTTTGATTATGCCACAGATTTCACAGCTGGATATGAGGCTC 747
Db 2145 ACCGATGGTGGCGCTTCGACTATGCCGGAGGTCGCGACAGTCGACTTTCAGCGCC 2204
Qy 748 TGAATAAAGATATCAAAAGATTGCTGACAACTTCCAGAGATTGTGTCCTCGCGATTATG 807
Db 2205 TGACCGCGGACATCGAGGAAGTGATGACCACTCGCAGCGGTGTGTGGCGCGGACTACG 2264
Qy 808 GTCAATTATGCTCTTTCTTTATGTTATGCTGTTGCTGCGACGGTTCGCGAACATACAGGACAT 867
Db 2265 GCCACTACGGGCGCTGTTTATCGGATGGGTGGCAGCTGCGCGACCTACCGCATCC 2324
Qy 868 ATGATGCGCGGGAGGCGCAGTGGTGTGATGAGCAACGTTTGAACCGCTGAACAGCTGGC 927
Db 2325 ACAGCGCGCGGGCGCGCGGGCGGCATGACGCGGTTTCGCGCGCTTAAACAGCTGGC 2384
Qy 928 CGGATAACGTTAATCTCGATAAAGCCCGCTGATTTGCTGTGGCCAGTCAAGAAAAATACG 987
Db 2385 CCGACAACGCGAGCTTGGACAAGCGCGCGGCTGCTGTGGCGGTCAAGAGAAGTACG 2444
Qy 988 GCTCCAGTATTTCTGGGGAGACTGATGTTCTGACTGGTAA--TGTTGCCCTTGAATC 1045
Db 2445 GCAAGAAGCTCTCATGGCGGACCTGATGTTGTTTCGCGCGGACCGCTGCGCTCGGAATC 2504
Qy 1046 CATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAAGTACCTGGGAGTCGGACCT 1105
Db 2505 GATGGGCTTCAAGACGTTTCGGTTTCGGTTCGG--GGTCGACAGTGGGAGCCGATGA 2562
Qy 1106 GGTATACCTGGGGGCTTCACAAACAGCCCTCTTTCAGATAAACCGGGATAAAAACGGGAACCT 1165
Db 2563 GGTCTATTGGGGCAAGGAAGCCACCTGGCTCGGCGATGACGGTTTACAGCGTAAAGCGATCT 2622
```

```
Qy 1166 TCAGAAACCTCTTGGCGCCACGACGATGGGACTTATTTATGTCAATCCTGAAGCCCGG 1225
Db 2623 GGAGAACCCGCTGCGCGCGGTGCAGATGGGCTGATCTAGTGAACCGGAGGCGCGAA 2682
Qy 1226 TGAATAACGATCTCTGCTGCTCCGCGAAAGATATCAGGAAGCTTTTTCACGATGGC 1285
Db 2683 CGGCAACCCGCGACCCATGCGCGCGGTGCGACATTCGCGAGACGTTTCGGGCGCATGGC 2742
Qy 1286 CATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATACATTTGTTAAAGCACA 1345
Db 2743 CATGAACGAGCTGGAACAGCGCGCTGATCGTCGGCGGTCAACATTCGTTAGAGCCCA 2802
Qy 1346 TGGTGCAGGCTCTCTCTGAAAAATGTATTGGCGCAGGCGCTGTATGGTGACCTGTGAGGA 1405
Db 2803 TGGGCGCG--GCCGCGCGATCTGCTCGGCCCGGAACCGGAGGCTGCTCGCTGGAGCA 2859
Qy 1406 GCAGGAGCTGGGATGGAATAAATAATGTGTGTGTACAGGAACCGCAATATACCATCACCAG 1465
Db 2860 GATGGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGGTAAGGACGCGATCACCG 2919
Qy 1466 TGGCCTGGAGGAGCCTGCTCGA 1488
Db 2920 CGCATCGAGGTCGTATGGACGA 2942

RESULT 17
US-08-313-185-45
; Sequence 45, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-313-185-45
```

CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention
 XX
 SQ Sequence 215 BP; 42 A; 63 C; 58 G; 52 T; 0 U; 0 Other;
 Query Match 67.7%; Score 17.6; DB 4; Length 215;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 3 GGGGTTCCAAAGCCCACTGACGA 26
 Db 111 GGTGTTCCGAGCCGCACTGACCA 88
 RESULT 13
 ABS23445/c
 ID ABS23445 standard; DNA; 215 BP.
 AC ABS23445;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe ORF from lung SEQ ID No 23436.
 XX
 KW Human; de; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenar syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 4; SEQ ID NO 23436; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC ; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)

CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types comprising one
 CC of the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagenar syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe open reading frame of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 215 BP; 42 A; 63 C; 58 G; 52 T; 0 U; 0 Other;
 Query Match 67.7%; Score 17.6; DB 6; Length 215;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 3 GGGGTTCCAAAGCCCACTGACGA 26
 Db 111 GGTGTTCCGAGCCGCACTGACCA 88
 RESULT 14
 AAK11215/c
 ID AAK11215 standard; DNA; 513 BP.
 XX
 AC AAK11215;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 11206.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000667.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
XX Example 4; SEQ ID NO 11206; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
XX Sequence 513 BP; 108 A; 152 C; 135 G; 118 T; 0 U; 0 Other;
SQ
Query Match 67.7%; Score 17.6; DB 4; Length 513;
Best Local Similarity 83.3%; Pred. No. 3e+02; Mismatches 0; Gaps 0;
Matches 20; Conservative 0; Indels 4; Indels 0; Gaps 0;
QY 3 GGGGTTCCAGCGCGCACTGACCA 26
DB 204 GGTGTTCCGAGCGCCACTGACCA 181
RESULT 15
ID ABS11018/c
XX
XX ABS11018;
XX
XX 19-AUG-2002 (first entry)
XX Human genome-derived single exon probe from lung SEQ ID No 11009.
XX Human; db; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX Homo sapiens.
XX
XX W0200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX Claim 1; SEQ ID NO 11009; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 513 BP; 108 A; 152 C; 135 G; 118 T; 0 U; 0 Other;
SQ
Query Match 67.7%; Score 17.6; DB 6; Length 513;
Best Local Similarity 83.3%; Pred. No. 3e+02; Mismatches 0; Gaps 0;
Matches 20; Conservative 0; Indels 4; Indels 0; Gaps 0;
QY 3 GGGGTTCCAGCGCGCACTGACCA 26
DB 204 GGTGTTCCGAGCGCCACTGACCA 181
RESULT 16
ID ACF57124 standard; cDNA; 787 BP.
XX
XX ACF57124;
XX
XX 14-OCT-2003 (first entry)
XX Human fatty acid CoA ligase-like AMP-binding enzyme EST SEQ ID NO:6.
XX Human; fatty acid CoA ligase-like AMP-binding enzyme; enzyme; neotropic;
KW neuroprotective; analgesic; antiparkinsonian; haemostatic; antianaemic;
KW immunostimulant; anorectic; antidiabetic; gene therapy; obesity; anaemia;
KW diabetes; haematological disorder; central nervous system disorder;
KW neuropathia; thrombocytopaenia; Alzheimer's disease; pain; EST;
KW Parkinson's disease; expressed sequence tag; chromosome 12; 12q13.13;
KW gene; ss.
XX
XX Homo sapiens.
XX
XX W02003057867-A2.
XX
XX 17-JUL-2003.
XX

PF 09-JAN-2003; 2003WO-EP000140.
XX
PR 10-JAN-2002; 2002US-0346602P.
PR 23-AUG-2002; 2002US-0405289P.
XX
PA (FARB) BAYER AG.
XX
PI Xiao Y;
XX
DR WPI; 2003-587123/55.
XX
PT New genes and its encoded fatty acid CoA ligase-like AMP-binding enzyme
PT (FACLAMP-BE), useful for identifying modulators of FACLAMP-BE activity,
PT and in gene therapy for treating e.g. obesity, diabetes, anemia,
PT Alzheimer's or pain.
XX
PS Disclosure; Page 115; 127pp; English.
XX
CC The present invention describes a human fatty acid CoA ligase-like AMP-
CC binding enzyme (I). (I) has neurotropic, neuroprotective, analgesic,
CC antiparkinsonian, haemostatic, antianaemic, immunostimulant, anorectic
CC and antidiabetic, and can be used in gene therapy. (I) polynucleotides
CC and polypeptides can be used for identifying test compounds, that may act
CC as agonists or antagonists at the receptor site and which can be
CC regulated to provide therapeutic effects. Vectors comprising the
CC polynucleotide can be used in the preparation of a medicament for
CC modulating the activity of (I) in a disease, particularly obesity,
CC diabetes, a haematological disorder or a central nervous system (CNS)
CC disorder. In particular, these are useful for treating anaemia,
CC neutropaenia, thrombocytopaenia, Alzheimer's disease, Parkinson's disease
CC and pain. The present sequence represents an expressed sequence tag (EST)
CC related to (I), which is given in the exemplification of the present
CC invention
XX
SQ Sequence 787 BP; 190 A; 193 C; 215 G; 189 T; 0 U; 0 Other;
Query Match 67.7%; Score 17.6; DB 9; Length 787;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 3 GGGGTTCCAGCGCCCACTGACGA 26
Db 638 GGAGCTCCAGCGCCCACTGAGA 661
RESULT 17
AAC98104
ID AAC98104 standard; cDNA; 3731 BP.
XX
AC AAC98104;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:114.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200055351-A1.
PN
XX 21-SEP-2000.
PD
XX 08-MAR-2000; 2000WO-US005883.
XX
XX 12-MAR-1999; 99US-0124270P.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587534/55.
DR P-PSDB; AAB53347.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.
XX
PS Claim 1; Page 552-553; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 3731 BP; 938 A; 914 C; 902 G; 971 T; 0 U; 6 Other;
Query Match 67.7%; Score 17.6; DB 3; Length 3731;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 3 GGGGTTCCAGCGCCCACTGACGA 26
Db 396 GGGTTCCTCCAGCGCCCACTGACCA 419
RESULT 18
ADQ67377
ID ADQ67377 standard; cDNA; 4066 BP.
XX
AC ADQ67377;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human cDNA sequence #2350.
XX
KW ss; gene; osteopathic; neuroprotective; neurotropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
OS Homo sapiens.
XX
XX EPI440981-A2.
PN
XX 28-JUL-2004.
PD
XX 21-JAN-2004; 2004EP-00001196.
PF
XX 21-JAN-2003; 2003JP-00102206.
PR
XX 09-MAY-2003; 2003JP-00131392.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
XX WPI; 2004-535376/52.
DR
DR P-PSDB; ADQ67684.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX
PS Claim 1; SEQ ID NO 4538; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 4066 BP; 992 A; 986 C; 959 G; 1129 T; 0 U; 0 Other;
Query Match 67.7%; Score 17.6; DB 12; Length 4066;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 GGGGTTCCAAAGCCGCAACTGACGA 26
||| ||||| ||||| ||||| |||||
DB 760 GGTGTTCCGAGCGCCACTGACCA 783
RESULT 19
ABL16366
ID ABL16366 standard; DNA; 4324 BP.
XX
AC ABL16366;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 571.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 571; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 4324 BP; 1148 A; 951 C; 1076 G; 1149 T; 0 U; 0 Other;
Query Match 67.7%; Score 17.6; DB 4; Length 4324;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAGGGGTTCCAAAGCCGCAACTGAC 24
||||| ||||| ||||| ||||| |||||
DB 1605 AAGGGGTTCCAAATCAGGAATGAC 1628
RESULT 20
ACA57512
ID ACA57512 standard; cDNA; 8586 BP.
XX
AC ACA57512;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human adipocyte Selected Interacting domain, SID, cDNA #599.
XX
KW Human; ss; gene; prey; adipocyte; SID; selected interacting domain;
KW anorectic; antidiabetic; protein-protein interaction; diabetes;
KW yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS Homo sapiens.
XX
PN WO200286122-A2.
XX
PD 31-OCT-2002.
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Daviet L;
XX
XX WPI; 2003-103412/09.
DR P-PSDB; ABU70968.
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
PS Claim 7; Page 301-305; 382pp; English.
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence encodes a SID
CC (prey) protein of the invention

XX	Sequence	8586 BP; 2314 A; 2181 C; 2342 G; 1749 T; 0 U; 0 Other;	
XX	Query Match	67.7%; Score 17.6; DB 8; Length 8586;	
XX	Best Local Similarity	83.3%; Pred. No. 4.2e+02;	
XX	Matches	20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy	3	GGGTTTCCAAAGCCGCAACTGACGA 26	
Db	7062	GGTGTTCGAGCCGCACTGACCA 7085	
RESULT 21			
AAT85392			
ID	AAT85392	standard; cDNA; 8906 BP.	
XX	XX		
AC	AAT85392;		
XX	XX		
DT	27-MAR-1998	(first entry)	
XX	XX		
DE	Human TRIO phosphoprotein cDNA.		
XX	XX		
XX	TRIO; human; proliferative disease; neurodegenerative disease; oncogene;		
KW	cytoskeleton; actin; diagnosis; transgenic animal; anticense; therapy;		
KW	ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PH	Key	Location/Qualifiers	
FT	CDS	67..8652	
FT		/*tag= a	
XX	XX		
PN	W09735979-A1.		
XX	XX		
PD	02-OCT-1997.		
XX	XX		
PF	27-MAR-1997;	97WO-US005236.	
XX	XX		
PR	27-MAR-1996;	96US-0014214P.	
XX	XX		
PA	(DAND) DANA FARBER CANCER INST INC.		
XX	XX		
PI	Streuli M, Debant A, Serra-Pagea C;		
XX	XX		
WP	1997-489642/45.		
DR	P-PSDB; AAW27227.		
XX	XX		
PT	New nucleic acid encoding the TRIO phospho:protein - used for diagnosis		
PT	and treatment of proliferative and neurodegenerative diseases.		
XX	XX		
PS	Claim 8; Page 80-93; 141pp; English.		
XX	XX		
CC	This polynucleotide sequence encodes human TRIO (see AAW27227), a novel		
CC	phosphoprotein suggested to be a central organiser of multiple signaling		
CC	pathways, to be involved in the activation of oncogenes such as c-fos and		
CC	to induce transformation of cells. The sequence is a composite of		
CC	overlapping cDNA clones obtained by screening human WI-38 fibroblast cDNA		
CC	for interaction with leucocyte common antigen-related transmembrane		
CC	protein tyrosine phosphatase domains by an interaction trap assay. TRIO		
CC	mRNA was expressed in all 8 human tissues examined. Once isolated, TRIO		
CC	cDNA can be expressed in usual host/vector systems to produce recombinant		
CC	TRIO polypeptides, e.g. for antibody production or screening for		
CC	modulators of activity. TRIO nucleic acid probes can be used to		
CC	detect/quantify TRIO activity, particularly for diagnosis and phenotyping		
CC	of neoplastic or hyperplastic disease, or to identify cells at risk of		
CC	transformation. Antisense nucleic acids can be used to modulate TRIO		
CC	activity, to reorganise the actin cytoskeleton (claimed), e.g. in cases		
CC	of wound healing and/or tumour metastasis, to treat an oncogene		
CC	(claimed), or more generally to control growth, differentiation,		
CC	migration and/or survival of cells, e.g. regulation of the immune		
CC	response to infection, treatment of impaired immune response (as in		
CC	chronic granulomatous disease), control of apoptosis in cancer therapy,		
CC	and treatment of degenerative diseases (e.g. Parkinson's, Alzheimer's or		

```

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

Query Match      67.7%; Score 17.6; DB 5; Length 8906;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGGGTTCCAAAGCGCGCAACTGACGA 26
Db 7128 GGTGTTCCGAGCGCGCACTGACCA 7151

RESULT 23
ABA82699
ID ABA82699 standard; DNA; 8906 BP.
XX
AC ABA82699;
XX
DT 25-JAN-2002 (first entry)
XX
DE GTPase gene SEQ ID NO:85.
XX
KW Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
KW sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
KW anticense therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
KW osteomalacia; fibrous dysplasia; ds.
XX
OS Homo sapiens.
XX
PN WO20017327-A1.
XX
PD 18-OCT-2001.
XX
PF 21-JUN-2000; 2000WO-US016951.
XX
PR 05-APR-2000; 2000US-00543771.
XX
PS 03-APR-2000; 2000US-00544398.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Carulli JP, Little RD, Recker RR, Johnson ML;
XX
DR WPI; 2001-657171/75.
XX
KW New high bone mass (HBM) and Zmax1 genes and proteins useful for
PT modulating bone mass for the treatment of e.g. osteoporosis.
XX
PS Claim 79; Page 386-389; 443pp; English.
XX
CC The present invention describes the human Zmax1 gene and the high bone
CC mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
CC genes have osteopathic activities. The genes can be used in gene therapy,
CC antise therapy and in the production of vaccines. They can be used in
CC the diagnosis and treatment of bone disorders including osteoporosis,
CC Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
CC ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

Query Match      67.7%; Score 17.6; DB 5; Length 8906;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGGGTTCCAAAGCGCGCAACTGACGA 26
Db 7128 GGTGTTCCGAGCGCGCACTGACCA 7151

RESULT 24
ACC46019
ID ACC46019 standard; DNA; 8906 BP.

```

```

XX ACC46019;
XX
DT 02-JUN-2003 (first entry)
XX
DE Human GTPase gene TRIO.
XX
KW Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200292764-A2.
XX
PD 21-NOV-2002.
XX
PF 13-MAY-2002; 2002WO-US014876.
XX
PR 11-MAY-2001; 2001US-0290071P.
XX
PR 17-MAY-2001; 2001US-0291311P.
XX
PR 01-FEB-2002; 2002US-0353058P.
XX
PR 04-MAR-2002; 2002US-0361293P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PA (AMHP ) WYETH.
XX
PI Babiş P, Bex FJ, Yaworsky PJ, Bodine PV;
XX
DR WPI; 2003-129278/12.
XX
KW New transgenic animals (e.g. mice), useful as models for studying bone
PT density modulation, developing drugs for treating or preventing bone
PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
PT reduced bone density.
XX
PS Disclosure; Page 509-512; 603pp; English.
XX
CC The invention relates to novel transgenic animals expressing the high
CC bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
CC comprising an alteration of the gene encoding LRP5 or LRP6, or expressing
CC an LRP5 that is modulated by an altered gene control sequence introduced
CC by homologous or non-homologous recombination. The transgenic animals are
CC for the study of bone density modulation or bone mass modulation. The
CC invention has osteopathic and cytostatic activity. The polynucleotides of
CC the invention may have a use in gene therapy. The transgenic animals and
CC nucleic acids are for the study of bone density modulation, where the
CC bone mass is modulated relative to non-transgenic animals of the same
CC species in more than one parameter selected from bone density, bone
CC strength, trabecular number, bone size, or bone tissue connectivity. The
CC transgenic animals, nucleic acids and methods are useful for identifying
CC molecules involved in bone development, and for developing pharmaceutical
CC compositions, which may be employed for treating or preventing bone
CC diseases, e.g. osteoporosis, osteomalacia, rickets, Paget's disease, or
CC neoplasms of the bone. The transgenic animals and nucleic acids are also
CC useful in methods for diagnosing diseases involved in bone development,
CC or characterised by reduced bone density or mass. The present sequence is
CC used in the exemplification of the invention
XX
SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

Query Match      67.7%; Score 17.6; DB 8; Length 8906;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGGGTTCCAAAGCGCGCAACTGACGA 26
Db 7128 GGTGTTCCGAGCGCGCACTGACCA 7151

RESULT 25
ADB98713

```



```

PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2003-040583/03.
XX
XX New human secreted proteins encoded by genes contained in cDNA clones
XX (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
XX multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
XX West Nile fever.
XX
XX Disclosure; Page 1858-1860; 2423pp; English.
XX
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
XX encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
XX treating or ameliorating medical conditions e.g. by protein or gene
XX therapy. The genes are isolated from a range of human tissues disclosed
XX in the specification. The nucleic acids, proteins, antibodies and
XX (ant)agonists are useful in the diagnosis, treatment and prevention of:
XX (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections
XX
XX Sequence 6480 BP; 1351 A; 1590 C; 1858 G; 1681 T; 0 U; 0 Other;
SQ
Query Match 68.5%; Score 17.8; DB 10; Length 6480;
Best Local Similarity 90.5%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 GGGTTCACAGCGCGCAACTGAC 24
DB 880 GGGTTCACAGCGCGCACTGCC 860
RESULT 11
ADC86400/c
ID ADC86400 standard; DNA; 8400 BP.
XX
XX ADC86400;
XX
XX 01-JAN-2004 (first entry)
XX Human GPCR gene SEQ ID NO:853.
XX
XX da; gene; human; GPCR;
XX guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1270724-A2.
XX
XX 02-JAN-2003.
XX
XX 18-JUN-2002; 2002EP-00013517.
XX
XX 18-JUN-2001; 2001JP-00246789.
XX
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
XX P-PSDB; ADC86401.
XX

```

```

XX
XX New polynucleotide, useful for preparing a composition for treating a
XX patient in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor.
XX
XX Claim 1; SEQ ID NO 853; 28pp; English.
XX
XX The invention relates to a novel polynucleotide encoding a guanosine
XX triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX the invention may have a use in gene therapy. The polynucleotide and
XX polypeptide are useful for preparing a composition for treating a patient
XX in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor. The
XX polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
XX invention.
XX
XX Sequence 8400 BP; 1949 A; 1993 C; 1981 G; 2277 T; 0 U; 200 Other;
SQ
Query Match 68.5%; Score 17.8; DB 10; Length 8400;
Best Local Similarity 90.5%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 GGGTTCACAGCGCGCAACTGAC 24
DB 7903 GGGTTCACAGCGCGCACTGCC 7883
RESULT 12
AAK23878/c
ID AAK23878 standard; DNA; 215 BP.
XX
XX AAK23878;
XX
XX 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe SEQ ID NO: 23869.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 23869; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX

```

```
Query Match      18.4%; Score 274.2; DB 2; Length 4795;
Best Local Similarity 61.4%; Pred. No. 3.9e-87;
Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;

Qy 688 ATCCCTGGGGGCTGATTTGATATGATCCACAGATTTCAACAGCTGGATATGAGGCTC 747
Db 2145 ACCCGATGGGTGGCGCTTCGACTATGCGGGAGGTCCGACCACTGACACTTGACGCC 2204

Qy 748 TGAAGAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGGGATATG 807
Db 2205 TGACCGGGACATCGAGGAAGTATGACCACTTCGACCGGTGGTGGCCCGCGACTACG 2264

Qy 808 GTCAATTATGCTCTTTCTTTATTCGTATGCTTGGCAGCTGCGGAAACATACAGCAT 867
Db 2265 GCCACTACGGCGCTGTTTATTCGGATGGGTGACGCTGCCGACCTACCGCATCC 2324

Qy 868 ATGATGCGCGGGAGCGCCAGTGGTGTAGCAACGTTTTGAAACCGCTGAACAGCTGGC 927
Db 2325 ACAGCGCGCGCGCGCGCGCGCGCGCATGACGCGTTCGCGCGCTTAACAGCTGGC 2384

Qy 928 CGGATACGTTAATCTCGATAAAGCCGCTCGATTTGCTGTGCGCCAGTCAAGAAATAACG 987
Db 2385 CCGACAAGCCGAGCTTGGACAAGCGCGCGCGCTGCTGTGGCCGGTCAAGAAGATACG 2444

Qy 988 GCTCCAGTATTTCTGGGGAGACCTGATGGTCTCTGACTGGTAA--TGTTCCTTGAATC 1045
Db 2445 GCAAGAGCTCTCATGCGCGGACCTGATTTGTTTCGCGGCAACCGCTGCGCTCGGAATC 2504

Qy 1046 CATGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAATGACTGGGAGTGGACCT 1105
Db 2505 GATGGGCTTCAGACGTTCCGGTTCCGGTTCCG--GCGTCGACCACTGGGAGACCGATGA 2562

Qy 1106 GGTATACTGGGGGCTGACAAACAGCTCTTTCGAGATAACCGGGATAAAGCGGAACT 1165
Db 2563 GGTCTATTGGGGCAAGGAAGCCACTGCTGCTCGCGGATGACGCTTACAGCGTAAGCATCT 2622

Qy 1166 TCAGAAACCTCTTGCAGCAGCAGATGGGACTTATTTATGTCATCTGAAGCCCGG 1225
Db 2623 GGAGAACCCGCTGGCGCGCTGAGATGGGCTGATCTACGTGAACCCGAGGCGCGGAA 2682

Qy 1226 TGAAGAACCCAGATCTCTGGCTTCGCGAAAGATATCAGGGAAGCTTTTTCACGTATGC 1285
Db 2683 CGGCAACCCGACCCCATGGCGCGCGGCTCGACATTCGCGAGAGCTTTTCGCGCATGGC 2742

Qy 1286 CATGGATGATGAGAGACTGTGGCTGATCGCGGAGGATACATTTGTTGTTAAAGACA 1345
Db 2743 CATGAACGACGTCGAAACAGCGCGCTGATCGTCGCGGCTCACACTTTTCGTTAAGACCA 2802

Qy 1346 TGGTGCAGCGTCTCTGAAATAATGTTTGGCGCAGGCGCTCATGTGTCACCTGTGGAGGA 1405
Db 2803 TGGCGCGG---GCCCGCGGATCTGGTCCGCGCGGACCCGAGGCTGCTCCGCTGGAGCA 2859

Qy 1406 GCAGGGACTGGGATGGAATAAATGTTGTTACAGGAAACCGGCAATATACCATCACAG 1465
Db 2860 GATGGCTTGGGCTGGAAGAGCTGATGGCAACCGGAACCGTAAGGACGCGATCACCA 2919

Qy 1466 TGGCTGGAAGGACCTGGTGA 1488
Db 2920 CGGCATCGAGGTCTGATGGAGCA 2942
```

RESULT 18

```
US-09-082-614A-45
; Sequence 45, Application US/09082614A
; Patent No. 6124098
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Anallo
; APPLICANT: Bodmer, Thomas
```

```
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082.614A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/313,185
; FILING DATE: 12-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-082-614A-45
```

```
Query Match      18.4%; Score 274.2; DB 3; Length 4795;
Best Local Similarity 61.4%; Pred. No. 3.9e-87;
Matches 493; Conservative 0; Mismatches 303; Indels 7; Gaps 3;

Qy 688 ATCCCTGGGGGCTGATTTGATATGATCCACAGATTTCAACAGCTGGATATGAGGCTC 747
Db 2145 ACCCGATGGGTGGCGCGCTTCGACTATGCGCGGAGTCCGACCACTGACACTTGACGCC 2204

Qy 748 TGAAGAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGTCGCCCTGCGGATATG 807
Db 2205 TGACCGGGACATCGAGGAAGTATGACCACTTCGACCGCTGGTGGCCCGCGACTACG 2264

Qy 808 GTCAATTATGCTCTTTCTTTATTCGTATGCTTGGCAGCTGCGGAAACATACAGCAT 867
Db 2265 GCCACTACGGCGCTGTTTATTCGGATGGGTGACGCTGCCGACCTACCGCATCC 2324

Qy 868 ATGATGCGCGGGAGCGCCAGTGGTGTAGCAACGTTTTGAAACCGCTGAACAGCTGGC 927
Db 2325 ACAGCGCGCGCGCGCGCGCGCGCATGACGCGGTTTCGCGCGCTTAACAGCTGGC 2384

Qy 928 CGGATACGTTAATCTCGATAAAGCCGCTCGATTTGCTGTGCGCCAGTCAAGAAATAACG 987
Db 2385 CCGACAAGCCGAGCTTGGACAAGCGCGCGCGCTGCTGTGGCCGGTCAAGAAGATACG 2444

Qy 988 GCTCCAGTATTTCTGGGGAGACCTGATGGTCTCTGACTGGTAA--TGTTCCTTGAATC 1045
Db 2445 GCAAGAGCTCTCATGCGCGGACCTGATTTGTTTCGCGGCAACCGCTGCGCTCGGAATC 2504

Qy 1046 CATGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAATGACTGGGAGTGGACCT 1105
Db 2505 GATGGGCTTCAGACGTTCCGGTTCCGGTTCCG--GCGTCGACCACTGGGAGACCGATGA 2562

Qy 1106 GGTATACTGGGGGCTGACAAACAGCTCTTTCGAGATAACCGGGATAAAGCGGAACT 1165
```

Db	2563	GGTCTATTGGGGCAAGGAGCCACCTGGCTCGCGGATGACGGTTACAGCGTAAAGCGATCT	2622
Qy	1166	TCAGAAACCTTTTGGCCGCCACGAGATGGGACTTATTTATGTGTCATCTCTGAAGCCCGG	1225
Db	2623	GGAGAACCCGTGGCCGCGGTGCAGATGGGCGCTGATCTACGTGAACCCGCGAGCGCCGAA	2682
Qy	1226	TGGAAACAGATCCTCTGCTTCCGCGAAAGATATCAGGGAAGCTTTTTCAGTATGGC	1285
Db	2683	CGGCAACCCGAGCCCATGCGCGCGGTGCGCAATTCGCGAGACGTTTCGCGCATATGGC	2742
Qy	1286	CATGAGATGATGAGGAGACTGTGGCCCTGTGATCGCGGAGGGGCATACATTTGGTAAAGCAC	1345
Db	2743	CATGAACGAGCTGAAACAGCGGGCTGATCGTCGGCGGTACACATTTTCGGTAAGACCCA	2802
Qy	1346	TGTCGACGCTCTCTGAAATAATGATTTGGCGCAGGGCCGTGATGGTGACCTGTGGAGGA	1405
Db	2803	TGGCGCG--GCCCGCCGATCTGGTCGCGCCCGAACCCGAGGCTGCTCGCTGGAGCA	2859
Qy	1406	GCAGGACTGGATGGAAAAATAATGTGTGTACAGGAAACGGCAATATACCATCACAG	1465
Db	2860	GATGGGCTTGGCTGGAGAGCTCGTATGGCACCGGAACCGGTAAAGACGCGATCACAG	2919
Qy	1466	TGGCTCGAAGAGCCCTGGTCGA	1488
Db	2920	CGGCATCAGGTCGTATGGAGCA	2942
RESULT 19			
US-09-328-352-3989			
; Sequence 3989, Application US/09328352			
; Patent No. 6562958			
; GENERAL INFORMATION:			
; APPLICANT: Gary L. Breton et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER			
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: GTC99-03PA			
; CURRENT APPLICATION NUMBER: US/09/328,352			
; CURRENT FILING DATE: 1999-06-04			
; NUMBER OF SEQ ID NOS: 8252			
; SEQ ID NO 3989			
; LENGTH: 2181			
; TYPE: DNA			
; ORGANISM: Acinetobacter baumannii			
US-09-328-352-3989			
Query Match			
Beat Local Similarity 18.3%; Score 271.8; DB 4; Length 2181;			
Matches 478; Conservative 0; Mismatches 322; Indels 3; Gaps 1;			
Qy	684	TCAAATCCTCGGGGGCTGATTTTGTATATGCCACAGATTTCAACAGCTGGATATGGAG	743
Db	160	TCTAATCCACTCGATAAAGATTTTAAATPACCGTCAAGAATTTAAAAAACTCGACTACTAC	219
Qy	744	GCTCTGAAAAAGATATCAAGATTTGCTGACAACTCCAGGATTTGGTGCCCTGCGGAT	803
Db	220	GCACCTAAAGCCGATATCAAAAATGATTAACCGATTTCCCAAGACTGGTGGCCAGCCGAC	279
Qy	804	TATGCTCATATATGGTCTCTTTTATTCGTATGGCTTGGCACGGTTCGCGGAACATACAGG	863
Db	280	TGGGGAATTTACGGGCTTATTTATCGTTAGCGTGGCATGCTGCGCGTACATATCGT	339
Qy	864	ACATATGATGCGCGGGAGGCGCCAGTGGTGTGAGCAAGCTTTTGAACCCCTGACAGC	923
Db	340	ATGGCGGATGCGCGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	399
Qy	924	TGGCGGATAGTTAAATCTGATAAAGCCGTCGATTCGATTCGTCGCGGAGTCAAGAAAAA	983
Db	400	TGGCTTGACATGCAAGTTTATGATTAAGACGCCGCTTTACTATGCGCGAGTTAAACAAAA	459
Qy	984	TACGGCTCAGTATTTCTCGGGGAGACCTGATGGTCTGACTGGTAAATGTTGCCCTTGA	1043
Db	460	TACGGTCAAAAAATATCTTGGGCCGATTTATTTATTTCTTGGCGGAAACATCGCACTTGAG	519

Qy	1044	TCCATGGGATTTAAAAACGCTGGGATTTTGTGCGCGAAGAGAGATGATCTGGAGTCGGAC	1103
Db	520	TCCTTCGTTTCCGTACCTTTGGTTTGGTGCCTGCTGGAAGATGTTTGGGAACACAGAC	579
Qy	1104	CTGGTATATCTGGGGGCTGACAAAGCCTCTTTCGAGATAACCGGGATAAAAACGGGAAA	1163
Db	580	AACGATGTTAAACTGGGGCGATGAAAAAGATGGTTAGCTCATCGCAACTCTGAAGCATTTG	639
Qy	1164	CTTCAGAAACCTCTTCGCGCCACGAGATGGGACTTATTTATGTCAATCTCTGAAGGCCCC	1223
Db	640	GCGGCGAGTAATCTTGTGCAACCGAAATGGGCTAATTTATGTGAACCTCTGAAGGACCA	699
Qy	1224	GGTGGAAAAACAGATCCTCTGGGCTTCGCGAAAGATATCAGGGAAGCTTTTTCACGATG	1283
Db	700	CAAGCGAGTGGTGACCCACGATCAGCAGCACCGCTTTATTCGTGCCACTTTTCGGCAACATG	759
Qy	1284	GCCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGGCTATACATTTGGTAAAGCA	1343
Db	760	GCAATGGATGATGAAGAAATTTGTTGCCCTTGATTTGCAGCGGTCTACTCTTTAGGTAAACT	819
Qy	1344	CATGGTCAGCGCTCTCTGAAATGATTTGGCGCAGGCTGATGTGCACCTGTGGAG	1403
Db	820	CACGGCGCTCCCCCTGCTGACCA---TGTCAAGCAGACCCGGAAGGTGCACCAATTGAA	876
Qy	1404	GAGCAGGACTGGGATGAAAAATAAATGTGTGTCAGGAAACCGCAAAATATACCATCAC	1463
Db	877	CAATGGGCTTTGGTTGGGCAACAGCTATGTTACAGGTGTCGTTAAAGATGCGATCACT	936
Qy	1464	AGTGGCTGGAAGAGCGCTGGTC	1486
Db	937	TCAGGTTTAGAAGTATTATTGGTC	959
RESULT 20			
US-08-674-887A-5			
; Sequence 5, Application US/08674887A			
; Patent No. 5939300			
; GENERAL INFORMATION:			
; APPLICANT: Robertson, Dan E.			
; APPLICANT: Sanyal, Indrajit			
; APPLICANT: Adhikari, Robert S.			
; TITLE OF INVENTION: CATALASES			
; NUMBER OF SEQUENCES: 8			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Fish & Richardson P.C.			
; STREET: 4225 Executive Square, Suite 1400			
; CITY: La Jolla			
; STATE: CA			
; COUNTRY: US			
; ZIP: 92037			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: Windows95			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/674,887A			
; FILING DATE: 03-JUL-1996			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Haile, Ph.D., Lies A.			
; REGISTRATION NUMBER: 38,347			
; REFERENCE/DOCKET NUMBER: 09015/002001			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 619/678-5070			
; TELEFAX: 619/678-5099			
; INFORMATION FOR SEQ ID NO: 5:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2262 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			

```

XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
XX KW PCR primer; probe; ss.
XX OS Synthetic.
XX OS Escherichia coli.
XX PN W09955908-A2.
XX PD 04-NOV-1999.
XX XX
XX XX 27-APR-1999; 99WO-FR001000.
XX XX
XX XX 28-APR-1998; 98FR-00005329.
XX XX
XX XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX XX
XX XX Frechon DTM, Laure FC, Thierry D;
XX XX
XX XX WPI; 2000-013443/01.
XX XX
XX XX New nucleic acid containing sequences specific to enterohemorrhagic
XX PT Escherichia coli, particularly serotype O157:H7, used for detecting these
XX PT bacteria in food.
XX XX
XX PS Claim 5; Page 27; 48pp; French.
XX XX
XX XX AA236103-27 represent fragments derived from nucleic acid sequences
XX CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
XX CC derived from two sequences. The first (AA236101) is 99.9% homologous to
XX CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AA236101), and
XX CC 99.8% homologous with IS91 of E. coli (nucleotides 1-406 of AA236102).
XX CC The second sequence (AA236102) is associated with the presence of
XX CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
XX CC 237-570 of AA236102 also have 68% homology with the virK gene which codes
XX CC for virulence proteins of Shigella flexneri. Both sequences are of
XX CC plasmid origin. The fragments are used as PCR primers and probes for the
XX CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
XX CC in human or animal samples, foods or the environment. The fragments are
XX CC also useful for epidemiological studies
XX XX
XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;
XX
XX Query Match 80.8%; Score 21; DB 3; Length 31;
XX Best Local Similarity 100.0%; Pred. No. 5.3;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 6 GTTCCAGCGCGCAACTGACGA 26
DB 31 GTTCCAGCGCGCAACTGACGA 11
XX
RESULT 5
ACH87244
ID ACH87244 standard; DNA; 746 BP.
XX
XX ACH87244;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #20439.
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;
XX KW alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
XX
XX US2003194704-A1.
XX PN
XX 16-OCT-2003.
XX PD
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX PR (PENN/) PENN S G.
XX PA (RANK/) RANK D R.
XX PA (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
XX PT gene expression analysis, for identifying or characterizing alternative
XX PT splicing events, for assessing genomic alterations or as tools for
XX PT surveying tissues.
XX
XX PS Claim 1; SEQ ID NO 20439; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
XX CC expression, comprising any of the 27,400 fully defined nucleotide
XX CC sequences in the specification, or their complements or fragments, and
XX CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
XX CC fully defined in the specification. The probe is a single exon probe that
XX CC hybridises under high stringency conditions to a nucleic acid molecule
XX CC expressed in human cells or tissues. Also included are a spatially-
XX CC addressable set of single exon nucleic acid probes for measuring human
XX CC gene expression (comprising a plurality of single exon nucleic acid
XX CC probes cited above, where each of the plurality of probes is separately
XX CC and addressably isolatable or amplifiable from the plurality), a single
XX CC exon microarray for measuring human gene expression, a method of
XX CC measuring human gene expression, a vector comprising the single exon
XX CC probe cited above, an ORF-encoded peptide comprising at least 8
XX CC contiguous amino acids of any of the above-mentioned amino acid
XX CC sequences (optionally with conservative amino acid substitutions), an
XX CC isolated antibody that binds specifically to a peptide cited above,
XX CC methods of selling and/or licensing single exon probes or microarrays to
XX CC a customer desiring to measure gene expression, a method of providing
XX CC human gene expression data by subscription, and a computer-readable
XX CC storage medium which contains a database having a plurality of records
XX CC (each record including data on the expression of a single exon probe
XX CC cited above. The probe, methods and apparatus are useful in gene
XX CC expression analysis. The probes may be used as tools for surveying
XX CC tissues to detect the presence of expressed messages that contain their
XX CC specific exon, or in constructing genome-derived single exon microarrays.
XX CC In addition, the probes are used in identifying and characterising
XX CC alternative splicing events, in detecting and characterising gross
XX CC alterations in the genomic locus that includes their exon, in assessing
XX CC smaller genomic alterations, in priming the synthesis of nucleic acids,
XX CC or in expressing the ORF-encoded peptide. The present sequence is a human
XX CC single exon probe of the invention. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
XX SQ Sequence 746 BP; 176 A; 202 C; 232 G; 136 T; 0 U; 0 Other;
XX
XX Query Match 69.2%; Score 18; DB 12; Length 746;
XX Best Local Similarity 80.8%; Pred. No. 2e+02;
XX Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 1 AAGGGGTTCCAGCGCGCAACTGACGA 26
DB 668 AAGGGGTTCCAGCGCGCAACTGACGA 693
XX
RESULT 6
AA25196
ID AA25196 standard; cDNA; 1924 BP.
XX
XX AA25196;
XX AC
XX XX
XX DT 19-JUL-1999 (first entry)
XX

```


CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of *Shigella flexneri*. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 26 BP; 8 A; 7 C; 8 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.023;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGGGGTTCCAAAGCGCAACTGACGA 26

Db 1 AAGGGGTTCCAAAGCGCAACTGACGA 26

RESULT 2

AAZ36101/c
 ID AAZ36101 standard; DNA; 1489 BP.

XX AC AAZ36101;

XX DT 11-FEB-2000 (first entry)

XX DE Nucleic acid sequence specific to enterohemorrhagic *Escherichia coli*.

XX KW Enterohemorrhagic *Escherichia coli*; EHEC; katP gene; *E. coli* O157:H7;
 XX IS91; ds.

XX OS *Escherichia coli*.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
 XX *Escherichia coli*, particularly serotype O157:H7, used for detecting these
 XX bacteria in food.

XX PS Claim 1; Fig 1; 48pp; French.

XX CC The present sequence is specific to enterohemorrhagic *Escherichia coli*
 XX (EHEC). The sequence is 99.9% homologous to the katP gene of *E. coli*
 XX O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
 XX homologous with IS91 of *E. coli* (nucleotides 1-406 of the present
 XX sequence). The present sequence is of plasmid origin. Fragments of the
 XX present sequence are used, as probes and primers, for detection of *E.*
 XX *coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC), in human or
 XX animal samples, foods or the environment. The fragments are also useful
 XX for epidemiological studies

XX SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 3; Length 1489;
 Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGGGGTTCCAAAGCGCAACTGACGA 26

Db 426 AAGGGGTTCCAAAGCGCAACTGACGA 401

RESULT 3

AAZ36111

ID AAZ36111 standard; DNA; 26 BP.

XX AC AAZ36111;

XX DT 11-FEB-2000 (first entry)

XX DE Primer derived from a nucleic acid sequence specific to EHEC.

XX KW Enterohemorrhagic *Escherichia coli*; EHEC; katP gene; *E. coli* O157:H7;
 XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 XX PCR primer; probe; ss.

XX OS Synthetic.

XX OS *Escherichia coli*.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
 XX *Escherichia coli*, particularly serotype O157:H7, used for detecting these
 XX bacteria in food.

XX PS Claim 5; Page 27; 48pp; French.

XX CC AAZ36103-27 represent fragments derived from nucleic acid sequences
 XX specific to enterohemorrhagic *Escherichia coli* (EHEC). The fragments are
 XX derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 XX the katP gene of *E. coli* O157:H7 (nucleotides 407-1489 of AAZ36101), and
 XX 95.8% homologous with IS91 of *E. coli* (nucleotides 1-406 of AAZ36101).
 XX The second sequence (AAZ36102) is associated with the presence of
 XX virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 XX 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 XX for virulence proteins of *Shigella flexneri*. Both sequences are of
 XX plasmid origin. The fragments are used as PCR primers and probes for the
 XX detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC),
 XX in human or animal samples, foods or the environment. The fragments are
 XX also useful for epidemiological studies

XX SQ Sequence 26 BP; 7 A; 7 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 96.2%; Score 25; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.068;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGGGGTTCCAAAGCGCAACTGACG 25

Db 2 AAGGGGTTCCAAAGCGCAACTGACG 26

RESULT 4

AAZ36112/c

ID AAZ36112 standard; DNA; 31 BP.

XX AC AAZ36112;

XX DT 11-FEB-2000 (first entry)

XX DE Primer derived from a nucleic acid sequence specific to EHEC.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 38.0362 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-10
Perfect score: 26
Sequence: 1 aaggggtccaagcgcgaactgacga 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	26	3 AAZ36110	Aaz36110 Primer de
c 2	26	100.0	1489	3 AAZ36101	Aaz36101 Nucleic a
3	25	96.2	26	3 AAZ36111	Aaz36111 Primer de
c 4	21	80.8	31	3 AAZ36112	Aaz36112 Primer de
5	18	69.2	746	12 ACH87244	Ach87244 Human gen
6	18	69.2	1924	2 AAX25196	Aax25196 Maize 4-c
c 7	18	69.2	6578	10 ADF81808	Adf81808 Leukaemia
c 8	18	69.2	7315	4 AAK52613	Aak52613 Human pol
c 9	18	69.2	8291	4 AAK51629	Aak51629 Human pol
c 10	17.8	68.5	6480	10 ABZ67665	Abz67665 Human sec
c 11	17.8	68.5	8400	10 ADC86400	Adc86400 Human GPC
c 12	17.6	67.7	215	4 AAK23878	Aak23878 Human bra
c 13	17.6	67.7	215	6 ABS23445	Abs23445 Human gen
c 14	17.6	67.7	513	4 AAK11215	Aak11215 Human bra
c 15	17.6	67.7	513	6 ABS11018	Abs11018 Human gen
16	17.6	67.7	787	9 ACF57124	Acf57124 Human fat
17	17.6	67.7	3731	3 AAC98104	Aac98104 Human col
18	17.6	67.7	4066	12 ADQ67377	Adq67377 Novel hum
19	17.6	67.7	4324	4 ABL16366	Ab116366 Drosophil
20	17.6	67.7	8586	8 ACA57512	Aca57512 Human adi

21	17.6	67.7	8906	2 AAT85392	Aat85392 Human TRI
22	17.6	67.7	8906	5 AAS81211	Aas81211 DNA encod
23	17.6	67.7	8906	5 ABA82699	Abas82699 GTPase ge
24	17.6	67.7	8906	8 ACC46019	Acc46019 Human GTP
25	17.6	67.7	8906	10 ADB98713	Adb98713 Human GTP
26	17.6	67.7	8906	10 ADE82509	Ade82509 Human DNA
27	17.6	67.7	8906	10 ADD89083	Add89083 Encoding
28	17.6	67.7	8906	10 ADG14305	Adg14305 Human TRI
29	17.6	67.7	8908	5 AAS70499	Aas70499 DNA encod
30	17.6	67.7	9041	5 AAS84947	Aas84947 DNA encod
31	17.6	67.7	9437	5 AAS81212	Aas81212 DNA encod
32	17.6	67.7	9449	5 AAS70498	Aas70498 DNA encod
33	17.6	67.7	9449	8 ACD13399	Acd13399 Human DNA
34	17.6	67.7	9450	5 AAS84943	Aas84943 DNA encod
35	17.2	66.2	1191	6 AEN98279	Abn98279 Arabidops
36	17.2	66.2	1630	6 AAI19766	Aai19766 Human rib
37	17.2	66.2	2321	1 AAN91183	Aan91183 Spru gene
c 38	17.2	66.2	2339	12 ADQ23890	Adq23890 Human sof
39	17.2	66.2	4266	12 ADM67130	Adm67130 Human hom
c 40	17.2	66.2	110000	11 ADM27081_05	Continuation (6 of
c 41	17	65.4	31	3 AAZ36113	Aaz36113 Primer de
c 42	17	65.4	230	8 ABQ84722	Abq84722 Manduca s
c 43	17	65.4	360	3 AAC38390	Aac38390 Zea mays
c 44	17	65.4	628	3 AAC44073	Aac44073 Zea mays
c 45	17	65.4	879	8 ACA27238	Aca27238 Prokaryot

ALIGNMENTS

RESULT 1
AAZ36110
ID AAZ36110 standard; DNA; 26 BP.
XX
AC AAZ36110;
XX
DT 11-FEB-2000 (first entry)
XX
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
OS Synthetic.
OS Escherichia coli.
XX
FN WO9955908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
DR WPI; 2000-013443/01.
XX
PT New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 5; Page 27; 48pp; French.
XX
CC AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of


```
; SEQ ID NO 1
; LENGTH: 12839
; TYPE: DNA
; ORGANISM: TPRP-F1 GENOMIC CLONE
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5)..(11322)
; OTHER INFORMATION: "n"'s are any nucleic residue
US-09-125-287-1

Query Match      2.6%; Score 38.8; DB 3; Length 12839;
Best Local Similarity 43.8%; Pred. No. 0.15;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

Qy 424 CTTAGTATTTTGTCTGTAGTATCTATCCAGCAATAGGTATATCTGTTGCATCAATA 483
Db      |||||
Qy 3827 CTTATTGATGATTGTGGCAAGTATCCACCAAGACTTCAAGTGTCTTCAACCTAATC 3886
Db      |||||
Qy 484 AAGTTGACTTTTGTATACAAATCGCAATTTCCCTTAATCCGGAGCTATTTCGTATGATAA 543
Db      |||||
Qy 3887 AAGTTTCCACTTTTCTCACTTTCAATTATTACTCACTCTCAATTTATGTGTAC 3946
Db      |||||
Qy 544 AAAAAACTCTTCCTGTCTGATTCTTCTGGGGCTATCGGGGAGCTTTTCTACGGCTGTAG 603
Db      |||||
Qy 3947 AGTTGACATTTCAAGTATTAGGCCCAATTTTCTTAGCTCGGAATTTTTTAAATCTCTTT 4006
Db      |||||
Qy 604 CCGCTGATAAAAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACCTCTC 663
Db      |||||
Qy 4007 AAATATTTTGATTTATATCTACTTATTACGTAGTTTTTCAAGTATAAATTTTCATTA 4066
Db      |||||
Qy 664 TGAGATTACACAGCCCTGAATCAAAATCCCTGGGGGGCTGATTTTGTGATTATGCCACCAT 723
Db      |||||
Qy 4067 TATGAATTCACGGTCAAAATTTAAAGTTTATTATTAACCAATGCCACATTAACCTAGAACA 4126
Db      |||||
Qy 724 TTCAACAGCTGGATATGGAGCTCTGAATAAAGATATCAAGATTTGCTGACAACTTCCC 783
Db      |||||
Qy 4127 ACAATATCTTTTGTATCAAGATTGGCAATTCGTATACAAATATCTTTAGCAAGTAATA 4186
Db      |||||
Qy 784 AGGATTGGTGCCTCGCGATTATGGT 809
Db      |||||
Qy 4187 TGTATACCAACATTATGTAATATGAT 4212
```

Search completed: March 11, 2005, 13:12:32
Job time : 618.139 secs

```
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match      3.2%; Score 48; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 4.3e-05;
Matches 12; Conservative 217; Mismatches 157; Indels 0; Gaps 0;

Qy 1065 GGATTTCGCGGAAGAGAGACTGGGAGTCGGACCTGGTATACCTGGGGCCCTGAC 1124
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1445 GAATTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1386
Qy 1125 AACAGCCTCTGCAGATAACCGGATAAACAACGGAACTTCAGAAACCTCTTGC 1184
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1385 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1326
Qy 1185 ACGCAGATGGGACTTATTATGTCAATCCTGAAGCCCGGTGGAACACAGATCCT 1244
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1325 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1266
Qy 1245 GTTTCGCGAAGATATACGGAAGCTTTTTCAGTATGGCATGGATGATGAGAGACT 1304
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1265 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1206
Qy 1305 GTGCGCTGATCGCGGAGGCATACATTTGGTAAAGCATGTGTGCGAGCTCTCTG 1364
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1365 AAATGTATGGCGAGGCGCTGATGTGCACTGTGAGGAGCAGGAGCTGGGATGAAA 1424
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1425 AATAAATGTGTACAGGAACGGCAA 1450
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1085 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1060

RESULT 24
US-09-125-287-2

; Sequence 2, Application US/09125287B
; Patent No. 6114602
; GENERAL INFORMATION:
; APPLICANT: BARG, Rivka
; APPLICANT: SALTS, Tehiam
; TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAOPY IN
; FILE REFERENCE: INTR O GENETIC PARTHENOCAOPY IN PLANTS
; CURRENT APPLICATION NUMBER: US/09/125,287B
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: PCT/IL97/00051
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 4518
; TYPE: DNA
; ORGANISM: TPRP-F1 PROMOTOR
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (6)..(191)
; OTHER INFORMATION: "n"s are any nucleic residue
; US-09-125-287-2

Query Match      2.6%; Score 38.8; DB 3; Length 4518;
Best Local Similarity 43.8%; Pred. No. 0.062;
Matches 169; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

Qy 424 CTTAGTATTTTTCCTGCTAGTATCTATCCAGCAATAGGTATATCTCTGTCATCAATA 483
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 3826 CTTATTGATGATTTGGCAAGTATCCACCCAAAGACTTCAAGTGTCTTCAACCTAAATC 3885
Qy 484 AAGTTGACTTTTGTATACACATCGAATTTCCCTTAATCCGGAGCTATTCGTATGATAA 543
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 3886 AAGTTTCCACTTTTCTCACTTTCAATATTACTCTACTCTCAATTTATGTGGTAC 3945
Qy 544 AAAAACTCTTCTGTTCTGATTTCTTGGCGCTATCGGGAGCTTTTCTACCGCTGTAG 603
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 3946 AGTTGACATTTCAAGTATTAGGCCCAATTTTCTTAGCTCGGAATTTTAAATCTCTTT 4005
Qy 604 CCGCTGATAAAAGAGAGACTCAAAATTTCTACTATCCAGAAACACTGGATTTAACTCTC 663
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 4006 AAATATTTTGTATTATATCTACTTATTAGTGTAGTTTTCATAGTATAAATTTCAATTCATA 4065
Qy 664 TGAGATTACAGCCCTGAAATCAATCCCTGGGGGGCTGATTTTGTATGATTCACACAGAT 723
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 4066 TATGAATTCAGGTCAAAAATTTAAAGTTTATTTAACCAATGCCACATTAATTTAGAAAT 4125
Qy 724 TTCAACAGCTGGATATGAGGCTCTGAAAAAAGATATCAAAAGATTTGCTGACAACTTCCC 783
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 4126 ACAAATATCTTTTGTATCAAGATTTGGCAATTCGTATACAATAATCTTTAGCAAGTAATA 4185
Qy 784 AGATTTGTCCTGCGGATATAGGT 809
Db ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 4186 TGTATACCAACATTTATGTAATATGAT 4211

RESULT 25
US-09-125-287-1
; Sequence 1, Application US/09125287B
; Patent No. 6114602
; GENERAL INFORMATION:
; APPLICANT: BARG, Rivka
; APPLICANT: SALTS, Tehiam
; TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAOPY IN
; FILE REFERENCE: INTR O GENETIC PARTHENOCAOPY IN PLANTS
; CURRENT APPLICATION NUMBER: US/09/125,287B
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: PCT/IL97/00051
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
```


Db 2152486 AGATGGCTTGGCTGGAGAGCTCGTATGCGACCGAACCCTGAAGGACGCGATCACCA 2152427
QY 1465 GTGGCTTGGAGAGCGCTGTGCGA 1488
Db 2152426 GCGGCATCGAGGTCGTATGGACGA 2152403
RESULT 10
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R. M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1
Query Match 20.0%; Score 298.4; DB 3; Length 4411529;
Best Local Similarity 62.4%; Pred. No. 2.1e-93;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
QY 688 ATCCCTGGGGGCTGATTTGATTTATGATTCACACAGATTTCAACAGCTGGATATGAGGCTC 747
Db 2155943 ACCGATGGGTGGCGGCTTGCAGTATGCGCGGAGGTGCGGACCATCGACCTTGACGCCC 2155884
QY 748 TGA AAAAAGATATCAAGATTTGCTGACACTTCCAGGATGCTGCCCTGCGGATTTAG 807
Db 2155883 TGACGCGGACATCGAGGAAGTGTACCACTCGCAGCGCTGCTGGCGCCGCCACTACG 2155824
QY 808 GTCAATTATGCTTTCTTTATTCGTATGCTTGGCAGCTGCGGAGATACAGGACAT 867
Db 2155823 GCCACTACGGCCGCTGTTATTCGGATGGCGTGCGAGCTGCCGACCTACCGCATCC 2155764
QY 868 ATGATGGCGGGAGGCGCCAGTGTGTGAGCAACGTTTTGAAACCGCTGAACAGCTGGC 927
Db 2155763 ACAGCGCGCGCGCGCGCGCGCGCGCATGCGAGCGGTTGCGCGCGCTTAACAGCTGGC 2155704
QY 928 CGGATAACGTTAATCTGGATAAAGCCGTCGATGCTGTGGCCAGTCAAGAAATAACG 987
Db 2155703 CCGAACAACCGAGCTTGGACAAGGCGCGCGCGCTGCTGTGCGCGCTCAAGAAGAGTACG 2155644
QY 988 GCTCCAGTATTTCTGGGGAGACCTGTGCTGCTGACTGTAATGTTGCCCTTGAATCCA 1047
Db 2155643 GCAAGAAGCTCTCATGGCGGACCTTGATTTGTTCCCGGCGNACTGCGCGCTGGANATCGA 2155584
QY 1048 TGGGATTTAAACCGCTGGGATTTGCTGGCGGAAGAGAAGATGACTGGAGTCGGACCTGG 1107
Db 2155583 TGGGCTTCAAGACGTTTGGGTTGGCTTTCGGCGCGGTGACCAAGTGGAGCCCGATGAGG 2155524
QY 1108 TATAGTGGGGGCTGACAAACAGCCCTCTTCAGATTAACCGGGATA---AAGCGGAAC 1164
Db 2155523 TCTATTTGGGCAAGGAAGCAACCTCGCTCGCGCATGAGCGTTTACAGCGGTAAAGCGGGATC 2155464
QY 1165 TTCAGAACCTTCTCCGCGCAGATGGGACTTATTATCTCAATCTCTGAAGCCCGC 1224
Db 2155463 TGGAAACCCGCTGGCGCGGTGACAGATGGGCTGATCTAAGTGAACCGGAGGGCGCA 2155404
QY 1225 GTGGAACACCGATCTCTGGCTTCCGCGAAGATATCAGGAAGCTTTTTCACGTATGG 1284

Db 2155403 ACGGCAACCCGACCCCATATGGCGCGCGGTGCGACATTCGCGAGAGCTTTCGCGCATGG 2155344
QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATATACATTTGGTAAAGCAC 1344
Db 2155343 CCATGAACGACGTCGAAACAGCGCGCTGATCGTCGCGGTACACTTTCGGTAAGACCC 2155284
QY 1345 ATGGTCAGGCTCTCTCTGAAAAATGATTGGCGCAGGGCTGATGCTGCACCTGTGGAGG 1404
Db 2155283 ATGGCGCG---GCCCGCGCGATCTGCTCGGCGCCCAACCCGAGGCTGCTCCGCTGGAGC 2155227
QY 1405 AGCAGGAGCTGGGATCGAAAAATAAATGTGGTACAGAAACGGCAATATACCATCACCA 1464
Db 2155226 AGATGGCTTGGCTGGAAGAGCTGATATGGCAACCGGACCGGTAGGACCGCATCACCA 2155167
QY 1465 GTGGCTTGGAAAGGAGCTGTGTCGA 1488
Db 2155166 GCGGCATCGAGGTCGTATGGACGA 2155143
RESULT 11
US-09-489-039A-2501/c
; Sequence 2501, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2501
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2501
Query Match 19.1%; Score 284.8; DB 4; Length 1674;
Best Local Similarity 60.2%; Pred. No. 2.4e-91;
Matches 491; Conservative 0; Mismatches 322; Indels 3; Gaps 1;
QY 672 CACAGCCCTGAATCAAAATCCCTGGGGGGTGTATTTGATTTGCCACAGATTTCAACAG 731
Db 1665 CACTCCAATCGCTCCAAACCCGTTGGGTGAGAAATTTTAACTACAGGGAAGAGTTTAAAAAG 1606
QY 732 CTGATATGAGGCTCTGMAAAAGATATCAAGATTTGCTGACAACTTCCACGATGG 791
Db 1605 CTCGATTTACTCCGCGCTGAAGGCTGATCTCAGAGCGCTGCTGACCGATTCGCAAGATGG 1546
QY 792 TGCCCTCGGATTTATGTCATTATGTCCTTTTCTTTTATTCGTATGGCTTGGCACGGTGCC 851
Db 1545 TGGCGGCGAGCTGGGTAGCTATATTTGGCTGTTTATTCGTATGGCTGCGACGGCGCC 1486
QY 852 GGAACATACAGGACATATGATGGCGGGAGGCGCCAGTGTGTGTCAGCAACGTTTGA 911
Db 1485 GGCACCTACCGCAACGTTGACGCTGCGCGCGCGCGAGTCTGAGACAGCAGCGCTTTGCC 1426
QY 912 CCGCTGAACAGCTGGCGGGATAAGTTAATCTGGATAAGCCCGCTGATTCGTGTCGCA 971
Db 1425 CCCCTGAACCTCTGGCTGACACAGCTCAGCTCGATTAAGGCCCGCGCTGCTGCTGTCGCG 1366
QY 972 GTCAAGAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTCTCTGACTGGTAAT 1031
Db 1365 GTGAAACAGAAATATGGCAGAAATCTCTTGGGCCGACCTTTACATCTGCGCGGTAAC 1306
QY 1032 GTTGCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGAC 1091
Db 1305 GTCGCTCTGGAAGAACCGCGGCTTCGACACCTTTGGCTTTGGCGCGGTCGGAAGACGTC 1246
QY 1092 TGGAGTCTGAGCTGGTATATCTGGGGGCTGACAAACAGCTCTTTCAGATTAACCGGAT 1151

```

; MOLECULE TYPE: DNA
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: 70..2289
US-08-852-219-20

```

Query Match 20.0%; Score 298.4; DB 2; Length 2331;
Best Local Similarity 62.4%; Pred. No. 3.7e-96;
Matches 502; Conservative 0; Mismatches 296; Indels 6;

Qy	688	ATCCCTCGGGGGCTGATTTTGATATATGCGACCAAGATTTCAAACGCTGGATATGAGGGCTC	747
Db	236	ACCCGATTTGGGTGCGCGCTTCGACTATATGCGCGAGAGTTCGGACCATCGACGTTGACGCCC	295
Qy	748	TGAAAAAGAGATATCAAAAGATTGCTGACAACCTCCAGGATTTGTCGCCCTGCGGATATG	807
Db	296	TGACGCGGACATCGAGAAAGTGATGACCACTCGACGCCGTGTGTGGCCGCGGACTACG	355
Qy	808	GTCAATATGTCCTTTCTTTATTCGTATGGCTTGGCACGGTGC CGGAACATACAGGACAT	867
Db	356	GCCACTACGGGCGCGCTGTTTATCCGATGGCTGGCGACGCTGCCGCGACCTTACCGCATCC	415
Qy	868	ATGATGGCCGGGAGGCGCCAGTGGTGGTCAAGCAACGTTTTGAAACCGCTGAAAGCTGGC	927
Db	416	ACGACGCGCGCGCGCGCGCGCGGCGGCGCATGCGACGGGTTGCGCCGCTTAAACAGCTGGC	475
Qy	928	CGGATAACGTTAATCTCGATAAAGCCCGTCGATTGCTGTGCCAGTCAAGAAAAAATACG	987
Db	476	CCGACAACGCCAGCTTGGACAAAGCGCGCGCGGTGCTGTGGCCGCTCAAGAAAGATACG	535
Qy	988	GCTCCAGTATTTCTTGGGGAGACCTGATGTCCTGATCTGTTGCTTAATGTTGCCCTTGAATCCA	1047
Db	536	GCAAGAAGCTCTCATGGCGCGACCTGATGTGTTTTCGCGCGCAACTGCGCGCTGGAATCGA	595
Qy	1048	TGGGATTTAAAACGCTCGGATTTGCTGGCGGAGAGAGATGACTGGAGTCGGAACCTGG	1107
Db	596	TGGGCTTCAAGACGTTTCGGGTTCGGCTTCGGCCGGGTCGACCAAGTGGGAGCCCGATGAGG	655
Qy	1108	TATACTGGGGGCTGACAAACAGCCTCTTGCGAGATAACCGGGATA---AAAAACGGGAAC	1164
Db	656	TCATTGGGCAAGAGAACCACTGGCTCGGCGATGAGCGTTACAGCGGTAAAGCGGATC	715
Qy	1165	TTCAGAAAACCTTTGCGGCCACGAGATGGGACTTATTATGTCATCTGAAAGGCCCGC	1224
Db	716	TGAGAGAACCCGCTGGCGCGCGTGCAGATGGGGCTGATCTACGTGAACCCCGAGAGGGCGGA	775
Qy	1225	GTGGAACACAGATCCTCTGCGTTTCGCGGAAAGATATCAGGGAAGCTTTTTCGTTATGG	1284
Db	776	ACGGCAACCCGACCCCATGCGCGCGCGGTGCACATTCGCGAGACGTTTCGCGGCGATGG	835
Qy	1285	CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATACATTTGGTAAAGCAC	1344
Db	836	CCATGAACGACGTCGAAAACAGCGGCGCTGATCGTTCGCGCGTCAACATTTTCGGTAAAGCCC	895
Qy	1345	ATGGTCAGCGCTCTCTGAAAATGTATTTGGCGCAGGGCCCTGATGCTGACCTGTGGAGG	1404
Db	896	ATGGCGCG---GCCCGGCGCATCTGGTGGCGCCCGGAACCCGAGGCTGCTCCGCTGGAGC	952
Qy	1405	AGCAGGACCTGGGATGGAATAATAATGTGGTACAGGAAACCGCAATAACCATCACCA	1464
Db	953	AGATGGGCTTGGGCTGGAGAGCTCGTATGGCACCGGAACCGGTAAAGGACCGCATCACCA	1012
Qy	1465	GTGGCTGGAAGAGGCTGTGTGCA	1488
Db	1013	GCGGCATCGAGGTCGTATGGACCA	1036

RESULT 9

US-09-103-840A-2/c ; Sequence 2, Application US/09103840A
; Patent No. 6294328 ;
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.

```

; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Query Match	20.0%	Score 298.4	DB 3	Length 4403765
Best Local Similarity	62.4%	Pred. No. 2.1e-93		
Matches 502	Conservative 0	Mismatches 296	Indels 6	Gaps 2
Qy	688	ATCCTGGGGGGTGATTTTGGATTATGCCACACAGATTTCAACACAGCTGGATATCGAGGCTC	747	
Db	2153203	ACCGATGGGTGGGGTTGCACTATGCCCGGAGTTCGGACCATCGAGCTTGAACGCC	2153144	
Qy	748	TGAAAAAGATATCAAGATTTTCTGCACAACTTCCACGAGTTGGTGCCCTTCGCGGATTATG	807	
Db	2153143	TGACGGGGACATCGAGGAAGTGATACCACTCGCAGCCGTGGTGCCCGCCGACTACG	2153084	
Qy	808	GTCAATTATGTCCTTCTTTATTCGTATGGCTTGGCACCGTTCGCGAATACAGACAT	867	
Db	2153083	GCCACTACGGGCGCTGTTTATCCGATGGGTGGCACGTGCGGCACCTACCGCATCC	2153024	
Qy	868	ATGATGCGGGGAGCGCCAGTGGTGTCAGCAACGTTTTCGACCGCTGMAACAGCTGC	927	
Db	2153023	ACACCGCCCGCGCGCGCGGCGGCGATCAGCGGTTTCGGCCGCTTAACAGCTGC	2152964	
Qy	928	CGGATAACGTTAATCTGSGATAAAGCCCGCTCGATTGCTGGCCAGTCAAGAAAAATACG	987	
Db	2152963	CCGACAAACGACGCTTGGACAGGCGCGCGCTGCTGGCCGGTCAAGAAAGTACG	2152904	
Qy	988	GCTCGAGTATTTCTGGGAGACCTGATGGTCTGATCGTAAATGTTGGCTTGAATCCA	1047	
Db	2152903	GCAAGAAGCTCTCATGGGCGGACCTGATTGTTTTCGCGGCAACTCGCGGCTGGAATCGA	2152844	
Qy	1048	TGGGATTTAAACCGCTGGGATTTGCTGGCGGAAGAAAGATGACTGGGAGTCGGACCTCG	1107	
Db	2152843	TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCGGGTGCACAGTGGGAGCCGATGAGG	2152784	
Qy	1108	TATACTGGGGGCTGCACAAACAGCCTCTTCGACAGATAACCGGGATA---AAAAACGGGAAC	1164	
Db	2152783	TCTATTGGGCAAGGAAGCACCTTGCTCGCGGATGAGCGTTACAGCGGTAAAGCGGATC	2152724	
Qy	1165	TTCCAGAAACCTCTTTCGGCCACGCAGATGGGACTTATTTATGTCAATCTCTGAAGGCCCG	1224	
Db	2152723	TGAGAACCCGCTGGCGCGGTGCAGATGGGCTGATCTACGTGAACCCGAGAGGGCCGA	2152664	
Qy	1225	GTGAAAAACAGATCCTCTGGGCTTCGCGAAAAAGATATCAGGAAAGCTTTTTCACGCTATGG	1284	
Db	2152663	ACGGCAACCGGNACCCATGGCCGGCGGTGCACATTCGAGAGCTTTCGGCGCATGG	2152604	
Qy	1285	CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGGCATACATTTGGTAAAGCAC	1344	
Db	2152603	CATGAAACAGCTCGAAAACAGCGGCGCTGATCGTGGCGGTGCACACTTTCGGTAAAGCCC	2152544	
Qy	1345	ATGGTCACGGTCTCTGGAATAATGTATTGGCGAGGGCTGATGGTCACCTGTGGAGG	1404	
Db	2152543	ATGGCGCGG---GCCCGCGCGATCTGGGTTCGGCCCCGAAACCCGAGGCTGCTCCGCTGGAGC	2152487	
Qy	1405	AGCAGGACTGGGATGGAATAATAATGTGGTACAGGAACCGCAAAATATACATCACCA	1464	

```
QY 928 CGGATAACGTTAATCTGGATAAAGCCCTCGATTGCTGGCCAGTCAAGAAAAAATACG 987
  |||||
Db 416 CCGACACCCAGCTTGGACAAGCGGCCCGCTGCTGGCCGTCAAGAAGATACG 475
  |||||
QY 988 GCTCAGATATTCCTGGGAGACCTGATGCTCTGACTGTAATGTTGCCCTTGAATCCA 1047
  |||||
Db 476 GCAAGAAGCTCTCATGGCGGACCTGATTGTTTCGCGGCAACTGCGCGCTGGAATCGA 535
  |||||
QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAAAGATGACTGGGAGTCGGACCTGG 1107
  |||||
Db 536 TGGGCTTCAAGACGTTGGGTTTCGGCTTCGGCCGGGTGACCAAGTGGGAGCCGATGAGG 595
  |||||
QY 1108 TATACTGGGGCCCTGACAAACAGCCTCTTGCAGATAACCCGGATA---AAAACGGGAAC 1164
  |||||
Db 596 TCTATTGGGGCAAGGAAGCCACTTGGCTCGCGGATGACGTTACAGCGTGAAGCGGATC 655
  |||||
QY 1165 TTCAAGAAACCTCTTCCCGCCACGAGATGGGACTTATTATGTCAATCTTGAAGGCCCGC 1224
  |||||
Db 656 TGGAGAACCCCTGGCCGCGTGCAGATGGGCTGATCTACGTGAACCCGAGGGCCGA 715
  |||||
QY 1225 GTGAAAACACAGATCCTCTGGCTTCGGCCGGAAGATATCAGGGAAGCTTTTTCAGTATGG 1284
  |||||
Db 716 ACGGCAACCCCGGACCCATGCGCGCGGTGCGACATTCGCGGAGACGTTTCGGGCGATGG 775
  |||||
QY 1285 CCATGGATGATGAGGAGACTGTGCCCTGTATCGCGGAGGCGCATACATTTTGTAAAGCAC 1344
  |||||
Db 776 CCATGAACGCTCGAAACAGCGCGGTGATCTGCGCGGTGACACATTTTCGGTGAAGACC 835
  |||||
QY 1345 ATGGTGACGCTCTCTGAAAAATGTTATTTGGCGGAGGCGCTGATGTCACCTGTGGAGG 1404
  |||||
Db 836 ATGGCGCG---GCCCGCGCATCTGGTCGCGCCCGGAACCCGAGGCTGCTCCGCTGGAGC 892
  |||||
QY 1405 AGCAGGACTGGGATGGAATAATGTTGTGTCGCGAGGCGCTGATGTCAGCACTTACCATCA 1464
  |||||
Db 893 AGATGGCTTGGGCTTGAAGAGCTCGTATGGCAACCGGAACCGGTAAGGACGCGATCACA 952
  |||||
QY 1465 GTGGCCTTGAAGGAGCTGTGTCGA 1488
  |||||
Db 953 GCGGCATCGAGGCTGTATGGACGA 976
  |||||
```

RESULT 6

```
US-08-219-1
; Sequence 1, Application US/08852219
; Patent No. 5922575
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; STREET: 119 No. 5922575th Fourth Street, Ste. 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,219
; FILING DATE: 07-May-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandberg, Victoria A.
; REGISTRATION NUMBER: 41,287
; REFERENCE/DOCKET NUMBER: 230.00010130
; TELECOMMUNICATION INFORMATION:
```

RESULT 7

```
; TELEPHONE: 612-305-1226
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-852-219-1
```

```
Query Match 20.0%; Score 298.4; DB 2; Length 2235;
Best Local Similarity 62.4%; Pred. No. 3.6e-96;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
```

```
QY 688 ATCCCTGGGGGCTGATTTTGAATTATGCCACCAGATTTCAACAGCTGATATGAGGCTC 747
  |||||
Db 176 ACCCGATGGGTGGCGCTTCGACTATCCGCGGAGGTGCGGACCATCGAGTTGACGCC 235
  |||||
QY 748 TGAAGAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGTCCTCGGATATG 807
  |||||
Db 236 TGACGCGGACATCGAGGAAGTGTATGACCACTTCGACGCGTGTGGCCCGCGATACG 295
  |||||
QY 808 GTCAATTATGTCTTTCTTTATTCGTATGCTTGGCACGCTGCCGGAACATACAGGACAT 867
  |||||
Db 296 GCCACTACGGCCCTGTTTATCCGATGGCGTGGCACCTGCCGCGACCTTACCGCATCC 355
  |||||
QY 868 ATGATGCGCGGGAGGCGCCAGTGGTGTGAGCAACCGTTTTTGAACCGCTGAACAGCTGGC 927
  |||||
Db 356 ACGACGCGCGCGCGCGCGCGGCGCATGACGCGGTTTCGCGCGCTTAAACAGCTGGC 415
  |||||
QY 928 CGGATAACGTTAATCTGGATAAAGCCCTGATGCTGTGGCCAGTCAAGAAAAAATACG 987
  |||||
Db 416 CCGACAAACGCGAGCTTGGACAAGCGCGCGGCTGCTGTGGCGGTCAAGAAGAAGTACG 475
  |||||
QY 988 GCTCCAGTATTTCTGGGGAGACCTGATGCTGCTGACTGTTGTTGCTTGAATCCA 1047
  |||||
Db 476 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGCGCAACTGCGCGCTGGAATCGA 535
  |||||
QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAGATGACTGGGAGTCGGACCTGG 1107
  |||||
Db 536 TGGGCTTCAAGACGTTTCGGGTTCGGCTTCGCGCGGTGCGACCTGCGGAGCCCGATGAGG 595
  |||||
QY 1108 TATACTGGGGCTGACAAACAGCCTTTCGAGATAACCGGATA---AAAACGGGAAC 1164
  |||||
Db 596 TCTATTGGGGCAAGGAAGCCACTTGGCTCGCGGATGACGTTACAGCGTGAAGCGGATC 655
  |||||
QY 1165 TTCAAGAAACCTCTTCCCGCCACGAGATGGGACTTATTATGTCAATCTTGAAGGCCCGC 1224
  |||||
Db 656 TGGAGAACCCCTGGCCGCGGTGCGATGGGCTGATCTACGTGAACCCGAGGGCCGA 715
  |||||
QY 1225 GTGAAAACACAGATCCTCTGGCTTCGGCCGGAAGATATCAGGGAAGCTTTTTCAGTATGG 1284
  |||||
Db 716 ACGGCAACCCCGGACCCATGCGCGCGGTGCGACATTCGCGGAGACGTTTCGGGCGATGG 775
  |||||
QY 1285 CCATGGATGATGAGGAGACTGTGGCCCTGTATGTCGCGGAGGCGCATACATTTGTAAGCAC 1344
  |||||
Db 776 CCATGAACGCTCGAAACAGCGCGCTGATGTCGCGGTGACACATTTTCGGTGAAGACC 835
  |||||
QY 1345 ATGGTGACGCTCTCTGAAAAATGTTATTTGGCGGAGGCGCTGATGTCACCTGTGGAGG 1404
  |||||
Db 836 ATGGCGCG---GCCCGCGCATCTGGTCGCGCCCGGAACCCGAGGCTGCTCCGCTGGAGC 892
  |||||
QY 1405 AGCAGGACTGGGATGGAATAATGTTGTGTCGCGAGGCGCTGATGTCAGCACTTACCATCA 1464
  |||||
Db 893 AGATGGCTTGGGCTTGAAGAGCTCGTATGGCAACCGGAACCGGTAAGGACGCGATCACA 952
  |||||
QY 1465 GTGGCCTTGAAGGAGCTGTGTCGA 1488
  |||||
Db 953 GCGGCATCGAGGCTGTATGGACGA 976
  |||||
```


Db 542 TCTATTGGGCAAGGAACCCACTTGGCTCGCGCATGAGCGTTACAGCGGTAAGCGGATC 601
Qy 1165 TTCAAGAACTCTTGGCGCACGAGATGGACTTATTTATGCTCAATCTGAAGGCCCG 1224
Db 602 TGAGAAACCCGCTGGCGCGGTGAGATGGGCTGTACTAGTGAACCCGAGGGCCGA 661
Qy 1225 GTGGAACACAGATCTCTCTGGCTTCCCGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
Db 662 ACGCAACCCGACCCCATGGCGCGCGGTGCAATTCGAGACGCTTTCGGCGCATGG 721
Qy 1285 CCATGATGATGAGGACATGTGGCCCTGATCGCGGAGGCGCATACATTTGGTAAGCAC 1344
Db 722 CCATGAACGACGTGGAACACGCGCGCTGATCGTCGCGGTCAACATTCGGTAAGACCC 781
Qy 1345 ATGTGTCAGCTCTCTCTGAAATGTTATGCGCGAGGCGCTGTATGTTGTCACCTGTGAGG 1404
Db 782 ATGGCGCG--GCCCGCGCATCTGTGCGCCCGAACCGAGGCTGCTCCGCTGAGC 838
Qy 1405 AGCAGGACTGGGATGAAATAAATGTTGTACAGGAACGGCAATATACCATCACCA 1464
Db 839 AGATGGGCTTGGGCTGGAAGAGCTCGTATGCGACCGGAACCGGTAAGGACGCGATCACCA 898
Qy 1465 GTGGCTGGAGGAGCTGTGCGA 1488
Db 899 CGCGCATCGAGGTGCTATGGACGA 922

RESULT 19

ACA40570
ID ACA40570 standard; DNA; 2223 BP.

AC ACA40570;
DT 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #2227.
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX Mycobacterium tuberculosis.
XX WO20027183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU36700.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 28440; 1766bp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2223 BP; 428 A; 695 C; 741 G; 359 T; 0 U; 0 Other;

Query Match 20.0%; Score 298.4; DB 8; Length 2223;
Best Local Similarity 62.4%; Pred. No. 5.1e-84;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

Qy 688 ATCCCTGGGGGCTGATTTTGTATTATGCCACACAGATTTCAACAGCTGGATGAGGCTC 747
Db 167 ACCGATGGTGGCGGTTTCGATATGCCGGGAGGTGCGGACCATCGAGCTTGACGCC 226
Qy 748 TGAATAAAGATATCAAGATTTTCTGACAACTTCCAGGATTTGGTCCCTGCGGATATG 807
Db 227 TGACGGGGACATCGAGGAAGTGATGACACCTCGCAGCGGTGGTGGCCGCCACTACG 286
Qy 808 GTCATATGTCCTTTCTTTATTTGATGCTTGGCAGCGGTGGCGGACATACAGGACAT 867
Db 287 GCCACTACGGGCGCTGTTTATCCGATGGCGTGGCAGCGTGGCCGACCTACCGCATCC 346
Qy 868 ATGATGCGCGGGAGGCGCCAGTGGTGTGTCAGCAACGTTTGAACCGCTGAACAGCTGGC 927
Db 347 ACGACG 406
Qy 928 CGGATAACGTTAATCTGGATAAAGCCCGCTCGATTTGCTGTGCGCCAGTCAAGAAAAATACG 987
Db 407 CCGACAACGCCAGCTTGGACAAGCGCGCGCGCGCTGTGTGGCCGCTCAAGAAGAAGTACG 466
Qy 988 GCTCCAGTATTTCTGGGGAGACCTGATGCTCTGACTGTTAATGTTGCCCTTGAATCCA 1047
Db 467 GCAGAAGCTCTCATGGGCGGACCTGATTTGTTTCGCGCGCACTGGCGCGTGAATCGA 526
Qy 1048 TGGGATTTAAACCGCTGGGATTTTCTGGCGGAAGAGATGACTGGGAGTTCGAGACCTGG 1107
Db 527 TGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCCGGTTCGACCATGGGAGCCCGATGAGG 586
Qy 1108 TATCTGGGGCTTGACAACAAGCCTCTTGCGAGATAACCGGGATA---AAACGGGAAAC 1164
Db 587 TCTATTGGGCAAGGAAGCCACCTGGCTCGGCGCATGAGCGTTACAGCGGTAAAGCGGATC 646
Qy 1165 TTCAAAACCTCTTGGCGCCACGAGATGGACTTATTATGTCAATCTTGAAGGCCCG 1224
Db 647 TGGAGAACCCGCTGGCGCGGTGAGATGGGGCTGATCTACGTGAACCCGAGGGCGCGA 706
Qy 1225 GTGGAACACAGATCTCTGCTTTCGCGAAGAGATATCAGGGAGAGCTTTTTCAGCTATGG 1284
Db 707 ACGCAACCCGAGCCCATGTGCCCGCGCGGTGCGACATTCGCGAGACGTTTCGGGCGCATGG 766

963	Qy	CTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCTTGGGGAGACGTGATGGTCTGT	1022
436	Db	CTGTGGCCGATCAAGACGAAGTACGGCGCCAAAGCTGTCTGGGGCCGACCTCTATGTGCTG	495
1023	Qy	ACTCGTAATGTTCCTTGAATCCATCGGATTTAAACGCTGGGATTTGCTGGCGGAAGA	1082
496	Db	GTCGGCAACGTGCGCCCTGGAGTCGATGGGCTTTCAAGACCTTTGGCTTCGCGCGCGCCGC	555
1083	Qy	GAAGATGACTGGGAGTCCGACCTGGTATATCTGGGGGCTGCACAAACGCTCTTTGCAGAT	1142
556	Db	GCCGACAGTGGGAGCCCGAAGACTGTACTGGGGTCCGGAAGCAGCTGGCTGGAGCAC	615
1143	Qy	AACCGGATAAAAACG---GGAACCTTCAGAAACCTCTTCGCGCCACGACAGATGGGACTT	1199
616	Db	AAGCGCTACAGCGCGAGCGGTGAGCTGGACTCGCCCCCTGGCGCGGTCCAGATGGGCTGT	675
1200	Qy	ATTTATGTCAATCTGAAGGCCCGGTGGAAACACGATCCTCTGGCTTTCGGGAAAGAT	1259
676	Db	ATCTACGTCAATCCGGAAGGCCCGAAACGGCAACCCGACCGCTGGCTTTCGGCGCGCGAC	735
1260	Qy	ATCAGGGAAGCTTTTTCAGCTATGGCCATGGATGATGAGGAGACTGTGGCCCTGATCGCG	1319
736	Db	ATTGCGAGAGCTTTCGCGCGGATGGCCATGAATGACGAGAGACCGTCGCGCCCTGATCGCC	795
1320	Qy	GGAGGGATACATTTGGTTAAAGCACATGTGTGCACGCTCTCCTGAAAAATGTATTGGCGCA	1379
796	Db	GCGGCGCACACCTTCGCAAGGCTCAGG---CGCGGGCGACGCTCCCTCGTGGGGCTT	852
1380	Qy	GGGCTGTATGGTGCACTGTGTGAGGAGCAGGGACTGGGATGGAAAAATAAATGTGGTACA	1439
853	Db	GAGCCGGAAGCGCGCGGATCGAGGCCCGCAGGGCTTTTGGTGGGCGAGCAGCAGCGCACG	912
1440	Qy	GGAACGGCAAAATATACCATCAACGATGGGCTTGAAGAGCGCTGGTC	1486
913	Db	GGCAAGGGCCCCGACGCGCATCACCGCGCGGCCGGAAGTGTCTGGAC	959

RESULT 18

ACA38594

ID ACA38594 standard; DNA; 978 BP.

XX

AC ACA38594;

XX

DT 19-JUN-2003 (first entry)

XX

DE Prokaryotic essential

• • • • •

KW Antisense; d

KW drug design; gene.

[illegible]

US Mycobacterium bovis. vv

XX
DN
W0200277102-22

FN
WQZ00277183-AZ.
YY

XX
PD
03-OCT-2002

XX
53-100-100-2002.

21-MAR-2002: 2002W0-IIS009107

[illegible]

PR 21-MAR-2001: 2001US-

PR 06-SEP-2001; 2001

PR 25-OCT-2001; 2001US-0342923P.

for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 26464; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 5213 sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/publ/pct_sequences](http://wipo.int/pub/publ/pct_sequences)

Sequence 978 BP: 187 A: 296 C: 341 G: 154 T: 0 II: 0 Other:

20.0%: Score 298.4: DB 8: Length 978:

but Local Similarity 62.48: Pred. No. 3.1e-84:

Conservative	Mismatches	Indels	Gaps
502	0	296	6

688 ATCCCTGGGGGGCTGATTTTGATTATGCCACCAGATTTC AACAGCTGGGATATGGAGGCTC 747

100

12345678910111213141516171819202122232425262728293031323334353637383940414243444546474849505152535455565758596061626364656667686970717273747576777879808182838485868788899091929394959697989910010110210310410510610710810911011111211311411511611711811912012112212312412512612712812913013113213313413513613713813914014114214314414514614714814915015115215315415515615715815916016116216316416516616716816917017117217317417517617717817918018118218318418518618718818919019119219319419519619719819920020120220320420520620720820921021121221321421521621721821922022122222322422522622722822923023123223323423523623723823924024124224324424524624724824925025125225325425525625725825926026126226326426526626726826927027127227327427527627727827928028128228328428528628728828929029129229329429529629729829930030130230330430530630730830931031131231331431531631731831932032132232332432532632732832933033133233333433533633733833934034134234334434534634734834935035135235335435535635735835936036136236336436536636736836937037137237337437537637737837938038138238338438538638738838939039139239339439539639739839940040140240340440540640740840941041141241341441541641741841942042142242342442542642742842943043143243343443543643743843944044144244344444544644744844945045145245345445545645745845946046146246346446546646746846947047147247347447547647747847948048148248348448548648748848949049149249349449549649749849950050150250350450550650750850951051151251351451551651751851952052152252352452552652752852953053153253353453553653753853954054154254354454554654754854955055155255355455555655755855956056156256356456556656756856957057157257357457557657757857958058158258358458558658758858959059159259359459559659759859960060160260360460560660760860961061161261361461561661761861962062162262362462562662762862963063163263363463563663763863964064164264364464564664764864965065165265365465565665765865966066166266366466566666766866967067167267367467567667767867968068168268368468568668768868969069169269369469569669769869970070170270370470570670770870971071171271371471571671771871972072172272372472572672772872973073173273373473573673773873974074174274374474574674774874975075175275375475575675775875976076176276376476576676776876977077177277377477577677777877978078178278378478578678778878979079179279379479579679779879980080180280380480580680780880981081181281381481581681781881982082182282382482582682782882983083183283383483583683783883984084184284384484584684784884985085185285385485585685785885986086186286386486586686786886987087187287387487587687787887988088188288388488588688788888989089189289389489589689789889990090190290390490590690790890991091191291391491591691791891992092192292392492592692792892993093193293393493593693793893994094194294394494594694794894995095195295395495595695795895996096196296396496596696796896997097197297397497597697797897998098198298398498598698798898999099199299399499599699799899910001001100210031004100510061007100810091010101110121013101410151016101710181019102010211022102310241025102610271028102910301031103210331034103510361037103810391040104110421043104410451046104710481049105010511052105310541055105610571058105910601061106210631064106510661067106810691070107110721073107410751076107710781079108010811082108310841085108610871088108910901091109210931094109510961097109810991100110111021103110411051106110711081109111011111112111311141115111611171118111911201121112211231124112511261127112811291130113111321133113411351136113711381139114011411142114311441145114611471148114911501151115211531154115511561157115811591160116111621163116411651166116711681169117011711172117311741175117611771178117911801181118211831184118511861187118811891190119111921193119411951196119711981199120012011202120312041205120612071208120912101211121212131214121512161217121812191220122112221223122412251226122712281229123012311232123312341235123612371238123912401241124212431244124512461247124812491250125112521253125412551256125712581259126012611262126312641265126612671268126912701271127212731274127512761277127812791280128112821283128412851286128712881289129012911292129312941295129612971298129913001

748 TGAAAAAGATATCAAAGATTGCTGACAACTTCCAGGATTGGTGCCCTGCCGATTATG 807

182 TGA CGCGGACATCGAGGAAGTGATGACCACTCGCAGCCGTGGTGGCCGCGACTACG 241

[illegible]

808 GICATTAAGGTCCTTTCTTTATTCGATTCGCTTGGCACCGCTGGCCGATACATACAGGACAT 809

242 GCCACTACGGGCGGCTGTTATCCGGATGGGTGGCAGCTGCCGGCACCTACCGCATCC 301

[illegible]

888 A T G A T G G C C G G G A G G C C A G T G G T C A G C A A C G T T T G A A C C G C T G A A C A G C T G G C 327

302 ACGACGGCCGCGCGCGCGCGGCATGAGCGGTTCGCGCCTTAACAGCTGGC 361

928 CCGATTAACGTTAATCTGGATAAAGCCCGTCGATTGCTGTGGCCAGTCAAGAAAAATACG 987

[illegible]

362 CCGACAAACGCCAGCTTGGAACAAGCGCGCGCTGCTGGTCAAGGAAGAGTACG 421

988 GCTCCAGTATTTCTCCGGGGAGACCTTGATGGTCCTGACTGGTAAATGTTGCCCTTGAATCCA 104

Db 55583 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCCGGGTCCAGCCAGTGGGAGCCCGATGAGG 55524
Qy 1108 TATACCTGGGGCTCAGCAACAGCTCTTCGAGATAACCGGGATA---AAACGGGAAC 1164
Db 55523 TCTATTGGGGCAAGGAAGCCACCTGGCTCGGCGATGAGCGTTACAGCGGTAAAGCGGATC 55464
Qy 1165 TTCAGAACTCTTCCGCCAGCAGATGGGACTTATTATGTCAATCTGAAGSCCCCG 1224
Db 55463 TGGAGAACCCGCTGGCCGGGTGAGATGGGCTGATCTACGTGAACCGGAGGGGCCGA 55404
Qy 1225 GTGGAACACAGATCTCTCGCTTCGCGAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
Db 55403 ACGGCAACCCGAGCCCATGGCCGGCGGTGCGACATTCGCGAGACGTTTCGGCGCATGG 55344
Qy 1285 CCATGGATGATGAGGAGACTGTGGCCCTGATGCGGGAGGGGATACATTTGGTAAAGCAC 1344
Db 55343 CCATGAACGAGCTCGAAACAGCGGCGCTGATGTCGCGCGTCAACATTTTCGTAAAGCCC 55284
Qy 1345 ATGGTGCAGCTCTCTGAAAATGATTTGGCGCAGGCGCTGATGTCGTCACCTGTGGAGG 1404
Db 55283 ATGGCGCG---GCCCGCGATCTGGTCGCGCCCGAACCCGAGGCTCTCGCTGGAGC 55227
Qy 1405 AGCAGGACTGGGATGGAATAAATGTTGTTACAGGAACCGCAAAATATACCATCACCA 1464
Db 55226 AGATGGGCTTGGGCTGGAAGAGCTGATGTCACCGCAACCGGTAAGCAGCGATCACCA 55167
Qy 1465 GTGGCTGGAAGGAGCCTGTGCGA 1488
Db 55166 CGGCGATCGAGTCTGTATGGACGA 55143

RESULT 25
AAI99683_21/c
Continuation (22 of 44) of AAI99683 from base 2100001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments LOCUS AAI99683 Accession AAI99683
WP Fragment Name Begin End
WP AAI99683_00 1 110000
WP AAI99683_01 100001 210000
WP AAI99683_02 200001 310000
WP AAI99683_03 300001 410000
WP AAI99683_04 400001 510000
WP AAI99683_05 500001 610000
WP AAI99683_06 600001 710000
WP AAI99683_07 700001 810000
WP AAI99683_08 800001 910000
WP AAI99683_09 900001 1010000
WP AAI99683_10 1000001 1110000
WP AAI99683_11 1100001 1210000
WP AAI99683_12 1200001 1310000
WP AAI99683_13 1300001 1410000
WP AAI99683_14 1400001 1510000
WP AAI99683_15 1500001 1610000
WP AAI99683_16 1600001 1710000
WP AAI99683_17 1700001 1810000
WP AAI99683_18 1800001 1910000
WP AAI99683_19 1900001 2010000
WP AAI99683_20 2000001 2110000
WP AAI99683_21 2100001 2210000
WP AAI99683_22 2200001 2310000
WP AAI99683_23 2300001 2410000
WP AAI99683_24 2400001 2510000
WP AAI99683_25 2500001 2610000
WP AAI99683_26 2600001 2710000
WP AAI99683_27 2700001 2810000
WP AAI99683_28 2800001 2910000
WP AAI99683_29 2900001 3010000
WP AAI99683_30 3000001 3110000
WP AAI99683_31 3100001 3210000
WP AAI99683_32 3200001 3310000
WP AAI99683_33 3300001 3410000
WP AAI99683_34 3400001 3510000
WP AAI99683_35 3500001 3610000
WP AAI99683_36 3600001 3710000

WP AAI99683_37 3700001 3810000
WP AAI99683_38 3800001 3910000
WP AAI99683_39 3900001 4010000
WP AAI99683_40 4000001 4110000
WP AAI99683_41 4100001 4210000
WP AAI99683_42 4200001 4310000
WP AAI99683_43 4300001 4403765

Query Match 20.0%; Score 298.4; DB 4; Length 110000;
Best Local Similarity 62.4%; Pred. No. 5.2e-83;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
Qy 688 ATCCCTGGGGGCTGATTTCGATATGCCACAGATTTCAACAGCTGATATGGAGGCTC 747
Db 53203 ACCCGATGGGTGCGCGCTTCGACTATGCCGCGAGGTGCGACCATCATCGCTTGACGCC 53144
Qy 748 TGAATAAAGATATCAAGATTTGCTGCAACATCTCCAGGATTTGTCGCTCCGGAATTATG 807
Db 53143 TGACCGGGACATCGAGGAAGTATGACCACTTCGACCGCTGGTGGCCCGGACTAGC 53084
Qy 808 GTCATTATGTCCTTTCTTTATTCGTATGCTTGGCAGCTGCGGGAACATATACAGGACAT 867
Db 53083 GCCACTACGGCCGCTGTTTATCCGATGGCGTGGCAGCTGCCGACCTACCGCATCC 53024
Qy 868 ATGATGCCCGGGAGCGCCAGTGGTGTGAGCAACGTTTGAACCCGCTGAACAGCTGCC 927
Db 53023 ACGACGCGCGCGCGCGCGCGGCGCATGTCGCGCGGTTCGCGCGCTTAAACAGCTGCC 52964
Qy 928 CGGATAACGTTAAATCTGGATAAAGCCGTCGATTGCTGTGGCCAGTCAAGAAAAAATACG 987
Db 52963 CCGACAACGCCGCTTGGACAAAGCGCGCGCTGCTGTGGCCGCTCAAGAAAGATAGC 52904
Qy 988 GCTCCAGATATTTCTGGGGAGACCTGATGTCCTGACTGTGTAATGTTGTCCTTGAATCCA 1047
Db 52903 GCAAGAAGCTCTCATGCGCGACCTGATGTTTTCGCGCGCAACTGCGCGCTGGAATCGA 52844
Qy 1048 TGGGATTTAAACCTGCGATTTCTGCGCGAAGAGATGACTGGGAGTGGGAGGCTGG 1107
Db 52843 TGGGCTTCAAGACGTTCCGGGTTCCGGCTTCGCGCGCGGTGACCCAGTGGGAGGCTGAGG 52784
Qy 1108 TATACGCGGCGCTGACAAACAAAGCTCTTTCAGAGATAACCGGGATA---AAAACGGGAAC 1164
Db 52783 TCTATTGGGGCAAGGAAGCCACCTGGCTTCGCGGATGAGCGTTACAGCGGTAAAGCGGATC 52724
Qy 1165 TTCAGAAACCTTTGCGCGCCACGAGATGGGACTTATTTATGTCAATCTCTGAAGSCCCCG 1224
Db 52723 TGGAGAACCCGCTGGCCGCGGTGCGATGGGCTGATCTACGTGAACCGCGAGGGCGCA 52664
Qy 1225 GTGGAACACCGATCTCTGCTTCGCGAAGAGATATCAGGGAAGCTTTTTCAGCTATGG 1284
Db 52663 ACGCAACCCCGACCCATCGCGCGGTCGACATTCGCGAGACGTTTCGCGCGCATGG 52604
Qy 1285 CCATGGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGGCATACATTTGTTAAAGCAC 1344
Db 52603 CCATGAACGAGCTCGAAACAGCGCGCTGATTCGTCGGCGGTACACTTTCGTTAAGACC 52544
Qy 1345 ATGGTGCAGGCTCTCTGAAAAATGTTATGTCGCGAGGGCCCTGATGGTGACCTCTGTGGAG 1404
Db 52543 ATGGCGCG---GCCCGCGCATCTGCTCGGCCCGCAACCCGAGGCTGCTCCGCTGGAGC 52487
Qy 1405 ACGAGGAGCTGGGATGGAATAAATGTTGTCAGGAACGGCAAAATATATACCATCACCA 1464
Db 52486 AGATGGGCTTGGGCTGGAAGAGCTGTTATGGCACCGGAACCGGTAAAGACGCGATCACCA 52427
Qy 1465 GTGGCTGGAAGGAGCCTGTGCGA 1488
Db 52426 CGGCGATCGAGTCTGTATGGACGA 52403

Search completed: March 11, 2005, 04:19:44
Job time : 2185.3 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 603.139 Seconds
(without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagtcggagatgaag.....ctggaaggagcctggtcgac 1489

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	Length DB ID	Description	
1	357.6	24.0	2238 2	US-08-674-887A-7	Sequence 7, Appli
2	357.6	24.0	2238 3	US-08-951-844-7	Sequence 7, Appli
3	357.6	24.0	2238 3	US-08-412-347-7	Sequence 7, Appli
4	298.4	20.0	2235 1	US-08-418-782-1	Sequence 1, Appli
5	298.4	20.0	2235 2	US-08-228-662-1	Sequence 1, Appli
6	298.4	20.0	2235 2	US-08-852-219-1	Sequence 1, Appli
7	298.4	20.0	2331 1	US-08-418-782-20	Sequence 20, Appli
8	298.4	20.0	2331 2	US-08-852-219-20	Sequence 20, Appli
9	298.4	20.0	4403765 3	US-09-103-840A-2	Sequence 2, Appli
10	298.4	20.0	4411529 3	US-09-103-840A-1	Sequence 1, Appli
11	284.8	19.1	1674 4	US-09-489-039A-2501	Sequence 2501, Ap
12	284.8	19.1	2208 4	US-09-489-039A-2376	Sequence 2376, Ap
13	274.2	18.4	2221 1	US-08-418-782-2	Sequence 2, Appli
14	274.2	18.4	2221 1	US-08-228-662-2	Sequence 2, Appli
15	274.2	18.4	2221 2	US-08-852-219-2	Sequence 2, Appli
16	274.2	18.4	4794 2	US-08-459-499-8	Sequence 8, Appli
17	274.2	18.4	4795 2	US-08-313-185-45	Sequence 45, Appli
18	274.2	18.4	4795 3	US-09-082-614A-45	Sequence 45, Appli
19	271.8	18.3	2181 4	US-09-328-352-3989	Sequence 3989, Ap
20	224.6	15.1	2262 2	US-08-674-887A-5	Sequence 5, Appli
21	224.6	15.1	2262 3	US-08-951-844-5	Sequence 5, Appli
22	224.6	15.1	2262 3	US-09-412-347-5	Sequence 5, Appli
23	48	3.2	7218 1	US-08-232-463-14	Sequence 14, Appli
24	38.8	2.6	4518 3	US-09-125-287-2	Sequence 2, Appli
25	38.8	2.6	12839 3	US-09-125-287-1	Sequence 1, Appli
26	37.8	2.5	832 4	US-09-621-976-2813	Sequence 2813, Ap
27	37.6	2.5	505 4	US-09-621-976-15639	Sequence 15639, A

Sequence 14193, A
Sequence 1723, Ap
Sequence 3561, Ap
Sequence 12608, A
Sequence 118, App
Sequence 15921, A
Sequence 15913, A
Sequence 565, App
Sequence 35607, A
Sequence 46966, A
Sequence 3646, Ap
Sequence 13114, A
Sequence 12657, A
Sequence 147, App
Sequence 147, App
Sequence 147, App
Sequence 147, App

ALIGNMENTS

RESULT 1
US-08-674-887A-7
; Sequence 7, Application US/08674887A
; Patent No. 5939300
; GENERAL INFORMATION:
; APPLICANT: Robertson, Dan E.
; APPLICANT: Sanyal, Indrajit
; APPLICANT: Adhikari, Robert S.
; TITLE OF INVENTION: CATALASES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,887A
; FILING DATE: 03-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09015/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...2235
; US-08-674-887A-7

Query Match 24.0% Score 357.6; DB 2; Length 2238;
Best Local Similarity 66.7%; Pred. No. 1.3e-117;
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

```
QY 683 ATCAATCCCTGGGGCTGATTTTGTATATCCACCAGATTTCAACAGCTGTATATGA 742
DB 168 ATCGGACCCAAACGACCGCGATTTTGACTATGCGGAAGATTTAAGAAGCTAGATCTGGC 227
QY 743 GGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTCGGGA 802
DB 228 AGCGGTAAANAGACCTGCAGCGCTAATGACAGATTCACAGGACTGGTGGCCAGAGA 287
QY 803 TTATGTCATTTATGGTCTTTCTTTTATTCGTATGGCTTGGCAGCGTCCGGAACATACAG 862
DB 288 TTACGGTCATTTATGGCCCTTCTTTATACGCATGGCTGGCAGACGCGCCGACCTACCG 347
QY 863 GACATATGATGGCGGGAGCGCCAGTGTGTGTCAGCAAGCTTTTCAACCGCTGACAG 922
DB 348 TATCGGTGATGGCGGTGGTGGCTCCCGCTCACAGCGCTTCCGCTCTCAATAG 407
QY 923 CTGCGCGGATAAGCTTAATCTGGATAAAGCCCGTCGATTTGCTGGCCAGTCAAGAAAAA 982
DB 408 CTGCGCAGACAATGCCAATCTGGATAAAGCAGCTTGCTTCTTTGGCCCATCAACAAAA 467
QY 983 ATAGCGCTCCAGTATTTCTGGGAGACCTGATGGTCTGACTGGTAAATGTTGGCCCTTGA 1042
DB 468 ATAGCGTCGAAAAATCTCTGGCGGATCTAATGATACTACAGAAAACTAGCTCTGGA 527
QY 1043 ATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGAGATGACTGGGAGTC--- 1099
DB 528 AACTATGGGCTTTAAAACTTTTGGTTTTCAGSTGGCAGAGCATGATGTATGGGAGCTGGA 587
QY 1100 GGACCTGGTATACCTGATAGGAGACTGTGGCCCTGATCGCGGAGGCAATCATTTGG 1336
DB 588 AGAAGATGTATCTGGGGAGCAGAAACCGAATGGCTGGGAGACAGCGCTATGAAGTGA 647
QY 1157 CGGAAACTTTCAGAAACCTTCTGGCGGACGACAGATGGGACTTATTTATGTCATCTGA 1216
DB 648 CCGAGAGCTCGAAAAATCCCTGGGAGCGGTACAAATGGGACTCATCTATGTATTAACCCGA 707
QY 1217 AGGCCCCGGTGGAAAAACAGATCTCTGGCTTCGCGGAAGATATCAGGAGACTTTTTC 1276
DB 708 AGGACCAACCGCAAGCGCAGACCCCTATCGCTGCTGCGCGTGATATTCGTGAGACTTTGG 767
QY 1277 ACGTATGCCATGATGATAGGAGACTGTGGCCCTGATCGCGGAGGCAATCATTTGG 1336
DB 768 CCGAATGCAATGAATGACGAAGAAACCGTGGCTCTCATACGGGTGGACACACTTCGG 827
QY 1337 TAAAGCATGGTGCAGCGTCTCTGAAAAATGATTTGGCGCAGCGCTGATGTGCAAC 1396
DB 828 AAAAAACCATGGTGTCTGCCGATCGGAGAAATATGTGGCCGAGAGCGCTGCGCGCAGG 887
QY 1397 TGTGAGGAGCAGGACTGGGATGGAATAAATAATGTTGTTACAGAAACGCAATATAC 1456
DB 888 TATTGAAGAAATGAGCTCTGGGCTGAAAAACACCTTACGGCACCGCGACACGCTGGGATAC 947
QY 1457 CATCACCAGTGGCTGGAGGAGCTGTGCGA 1488
DB 948 CATCACCAGTGGACTAGAAGCGCTGGACCA 979
```

RESULT 2

US-08-951-844-7
; Sequence 7, Application US/08951844
; Patent No. 6074860

GENERAL INFORMATION:

; APPLICANT: Robertson et al.
; TITLE OF INVENTION: Catalases
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESS: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA

ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,844
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/674,887
FILING DATE: July 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Charles J. Heron
REGISTRATION NUMBER: 28,019
REFERENCE/DOCKET NUMBER: 331400-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 NUCLEOTIDES
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: DNA
US-08-951-844-7

Query Match 24.0%; Score 357.6; DB 3; Length 2238;
Best Local Similarity 66.7%; Pred. No. 1.3e-117;
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

```
QY 683 ATCAATCCCTGGGGCTGATTTTGTATATGCCACCAGATTTTCAACAGCTGTATATGA 742
DB 168 ATCGGACCCAAACGACCGCGATTTTGACTATGCGGAAGATTTAAGAAGCTAGATCTGGC 227
QY 743 GGCTCTGAAAAAGATATCAAGATTTTGTGTCGACAACTTCCAGGATTTGGTCCCTCGGGA 802
DB 228 AGCGGTAAANAGACCTGCAGCGCTAATGACAGATTCACAGGACTGGTGGCCAGAGA 287
QY 803 TTATGTCATTTATGGTCTTTCTTTTATTCGTATGGCTTGGCAGCGTCCGGAACATACAG 862
DB 288 TTACGGTCATTTATGGCCCTTCTTTATACGCATGGCTGGGACAGCGCGCAGCTACCG 347
QY 863 GACATATGATGGCGGGAGCGCCAGTGTGTGTCAGCAAGCTTTTGAACCGCTGAACAG 922
DB 348 TATCGGTGATGGCGGTGGTGGCTCCCGCTCACAGCGCTTCCGCGCTCTCAATAG 407
QY 923 CTGCGCGGATAACGTTAATCTGGATAAAGCCCGTCGATTTGCTGGCCAGTCAAGAAAAA 982
DB 408 CTGCGCAGACAATGCCAATCTGGATAAAGCAGCGTTCGCTTCTTTGGCCCATCAACAAAA 467
QY 983 ATAGCGCTCCAGTATTTCTGGGAGACCTGATGGTCTGACTGTGTAATGTTGCCCTTGA 1042
DB 468 ATAGCGTCGAAAAATCTCTGGCGGATCTAATGATACTACAGGAACGCTAGCTCTGGA 527
QY 1043 ATCCATGGGATTTAAACGCTGGGATTTGCTGGCGGAGAGAGATGACTGGGAGTC--- 1099
DB 528 AACTATGGGCTTTAAAACTTTTGGTTTTCAGAGTGGCAGAGCATGATGTGGGAGCTTGA 587
QY 1100 GGACCTGGTATACCTGATAGGAGCTTGAACAAAGCCTCTTTCAGATTAACCGGGATAAA---AA 1156
DB 588 AGAAGATGTATCTGGGGAGCAGAAACCGAATGGCTGGGAGACAAGCGCTATGAAGTGA 647
QY 1157 CGGAAACTTTCAGAAACCTTTCGCGGACGACAGATGGGACTTATTTATGTCATCTGGA 1216
DB 648 CCGAGAGCTCGAAAAATCCCTGGGAGCGGTACAAATGGGACTCATCTATGTATTAACCCGA 707
QY 1217 AGGCCCCGGTGGAAAAACAGATCTCTGGCTTCGCGGAAGATATCAGGAGACTTTTTC 1276
DB 708 AGGACCAACCGCAAGCGCAGACCCCTATCGCTGCTGCGCGTGATATTCGTGAGACTTTGG 767
```

Db 1245 TGGGAACCGGATCTCGACGTCGACTGGGCGGATGAGAAAGATGGCTGGCCACCGTCAT 1186
Qy 1152 AAAAAGCGGAAATCTCAGAAACCTCTTGGCGGACGACGATGGAGCTTATTTATGTCAAT 1211
Db 1185 CCGGAAAGCGCTCGCGGAAACAGGCGCATCGGCGGACCGGAAATGGTCTGATTTATGTAAC 1126
Qy 1212 CCGTGAAGGCGCGGTGGAAACACAGATCTCTGGCTTCGCGGAAAGATATCAGGGAAGCT 1271
Db 1125 CCGGAAAGTCCGAAAGCGGCGGAGCGCTGTGCGCGCGGCGGCGCATTCGCGCAC 1066
Qy 1272 TTTTCACTGATGCGCATGATGATGAGGAGCTGTGGCCCTGATCGCGGAGGCGATACA 1331
Db 1065 TTGCGCAATATGCGATGACGATGAAGAGATCGTCGCGCTGATCGCGCGGCGCATAG 1006
Qy 1332 TTGCTGAAGCATGTGTGACGCTCTCTGAAAAATGATTTGGCGCAGGCGCTGATGTT 1391
Db 1005 CTGGGCAAAACCGCGGCGCGCAGAAAC---CAGCCACGTGCGGCGCGGCGGAGCC 949
Qy 1392 GCACCTGTGGAGGAGCAGGAGCTGGGATGGAATAAATAATGTTACAGGAAACGCAAA 1451
Db 948 GCGCGCTGGAAGCGGAGGCTGTGGCTGGCACTCGAGTACGAGCGCGCGCGCGG 889
Qy 1452 TATACCATCACCAGTGGCTGGAAGGAGCCTGGTCG 1487
Db 888 GAGCGCATCACCTCCGGTCTGGAAGTGGTCTGGACG 853

RESULT 12
US-09-489-039A-2376
; Sequence 2376, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2376
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2376

Query Match 19.1%; Score 284.8; DB 4; Length 2208;
Best Local Similarity 59.5%; Pred. No. 3e-91;
Matches 500; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

Qy 648 CTGGATTTAACTCTCTGAGATTACACAGCCCTGAATCAATCCCTGGGGGCTGATTTT 707
Db 151 CTGCGTGTGATCTGTAAACAGCACCTCAATCGCTCCACCCGTGGGTGAGAAATTT 210
Qy 708 GATTATGCCACAGATTTTCAACAGCTGATGAGGCTCTGAAAAAGATATCAAGAT 767
Db 211 AACTCAGGGAAGAGTTTTAAAAAGCTCGATTACTCCGCGCTGAAGGCTGATCTCAGAGCG 270
Qy 768 TTGCTGACAACTTCCAGGATTTGCTCGGATTTATGTTTATGTTCTTTCTTTT 827
Db 271 CTGCTGACGATTTCGAAGAGTGGTGGCGGAGACTGGGGTAGCTATATGGCTGTGTT 330
Qy 828 ATTCTGATGGCTTGGCAGGTCGCGGAAACATACAGGACATATGATGGCGGGGAGCGCG 887
Db 331 ATTCTGATGGCTTGGCAGGTCGCGGACCTTACCACCGTTGACGCTGCGGCGCGCA 390
Qy 888 AGTGTGTGACCAACGTTTTGAAACGCTGGAACAGCTGCGCGGATTAATCTGAT 947
Db 391 GGTCTGTGACAGCAGCGCTTTTCCCGCCCTGAACTCTGCGCTGACAAAGCTCAGCCTCGAT 450
Qy 948 AAAGCCCGCTGATCTGTGGCGACTCAAGAAATAACGGCTCCAGTATTTCTTGGGA 1007

Db 451 AAGCCCGCCGCTCTGCTGTGGCCGCTGAAACAGAAATATGGCCAGAAATCTCTCTGGGCC 510
Qy 1008 GACCTGATGGTCTCTGACCTGGTAATTTTGGCCCTTGAATCATGGGATTTAAAAAGCTGGGA 1067
Db 511 GACCTTTACATGCTGGCGGTAACCGTCTCTGGAAACCGCGGCTTCGCAACCTTTGCG 570
Qy 1068 TTTGCTGGCGGAGAGAGATGACCTGGGAGTGGACCTGGTATATCTGGGGGCTGACAC 1127
Db 571 TTTGGCGCGGTGCGGAAGAGCTCTGGGAACCGGATCTCGAGCTCGACTGGGGCGATGAG 630
Qy 1128 AAGCCTTTGCAGATAACCGGGATAAAAACGGGAAACTTTCAGAAACCTCTTCCCGCCAG 1187
Db 631 AAAGAGTGGCTGGCCCAACCGTCTCCGGAAGCCTGGCGMAACACAGCCCATCGCGCCACC 690
Qy 1188 CAGATGGAGCTTATTTATGTCAATCTTGAAGCCCCGGTGGAAAAACAGATCTCTTGGCT 1247
Db 691 GAAATGGGTCTGATTTATGTAAACCGGAAAGTCCGAAACCGCGGAGCGGCTGTCC 750
Qy 1248 TCCGCGAAGATATCAGGGAAGCTTTTTCACGTATGGCCATGATGATGAGGAGACTGTG 1307
Db 751 GCGCGCGCGGCAATTCGCGCCACCTTCGGCAATATGGCGATGGAGTGAAGAGATCGTC 810
Qy 1308 GCCCTGATCGCGGAGGCGCATACATTTGTTAAAGCACATGTTGTCAGCGCTCTCTGAAAAA 1367
Db 811 GCGCTGATCGCGCGCGGCATACGCTGGGCAAAACCCACGCGCGCGCAGAAAC---CAGC 867
Qy 1368 TGTATTGGCGCAGGCGCTGATGGTGCACTGTGGAGGAGCAGGAGCTGGGATGGAATAAT 1427
Db 868 CACGTCGCGCGGAGCGCGGCGCGCTGGAAAGCGCAGGCTGTGGGCTGGCACTCC 927
Qy 1428 AAATCTGTACAGGAACGCGCAATATATACCATCAGCTGGCCTGGAGGAGCTGGCTG 1487
Db 928 AGCTACGCGAGCGCGCGGCGGAGCGCGCATACCTCCGCTCTGGAAGTGTCTCTGGACG 987

RESULT 13
US-08-418-782-2
; Sequence 2, Application US/08418782
; Patent No. 5658733
; GENERAL INFORMATION:
; APPLICANT: Cockerill, Franklin R.
; APPLICANT: Kline, Bruce C.
; APPLICANT: Uhl, James R.
; TITLE OF INVENTION: Detection of Isoniazid Resistant Strains
; TITLE OF INVENTION: of M. Tuberculosis
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,782
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.141US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2221 base pairs
; TYPE: nucleic acid

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 10.5316 Seconds
(without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-10
Perfect score: 26
Sequence: 1 aaggggtccaagccgaactgacga 26

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.6	67.7	8865	4	US-09-949-016-4238
2	17.6	67.7	8906	2	Sequence 4238, Ap
3	17.6	67.7	239527	4	Sequence 1, Appli
4	17.2	66.2	1430	6	Sequence 15980, A
5	17.2	66.2	1430	6	Patent No. 5514590-1
6	17.6	67.7	8906	2	Patent No. 5514590
7	17.6	67.7	8906	2	Sequence 68, Appl
8	17.6	67.7	8906	2	Sequence 2320, Ap
9	17.6	67.7	8906	2	Sequence 2596, Ap
10	17.6	67.7	8906	2	Sequence 2514, Ap
11	17.6	67.7	8906	2	Sequence 11, Appl
12	17.6	67.7	8906	2	Sequence 9, Appli
13	17.6	67.7	8906	2	Sequence 1185, Ap
14	17.6	67.7	8906	2	Sequence 13727, A
15	17.6	67.7	8906	2	Sequence 13728, A
16	17.6	67.7	8906	2	Sequence 13729, A
17	17.6	67.7	8906	2	Sequence 352, App
18	17.6	67.7	8906	2	Sequence 16329, A
19	17.6	67.7	8906	2	Sequence 2, Appli
20	17.6	67.7	8906	2	Sequence 1, Appli
21	17.6	67.7	8906	2	Sequence 1, Appli
22	17.6	67.7	8906	2	Sequence 709, App
23	17.6	67.7	8906	2	Sequence 71, Appl
24	17.6	67.7	8906	2	Sequence 976, App
25	17.6	67.7	8906	2	Sequence 976, App
26	17.6	67.7	8906	2	Sequence 976, App
27	17.6	67.7	8906	2	Sequence 976, App

28 16.2 62.3 42574 4 US-09-949-016-17525 Sequence 17525, A
29 16.2 62.3 269223 4 US-09-596-002-41 Sequence 41, Appl
30 16 61.5 30 3 US-09-242-690A-25 Sequence 25, Appl
31 16 61.5 30 4 US-09-908-855-25 Sequence 25, Appl
32 16 61.5 462 2 US-08-743-637B-10 Sequence 10, Appl
33 16 61.5 462 3 US-08-526-840B-10 Sequence 10, Appl
34 16 61.5 601 4 US-09-949-016-72780 Sequence 72780, A
35 16 61.5 601 4 US-09-949-016-72781 Sequence 72781, A
36 16 61.5 1032 4 US-09-252-991A-9669 Sequence 9669, Ap
37 16 61.5 1678 4 US-09-673-395A-111 Sequence 111, App
38 16 61.5 2037 4 US-09-902-540-6164 Sequence 6164, Ap
39 16 61.5 2039 4 US-09-902-540-304 Sequence 304, App
40 16 61.5 2511 4 US-09-252-991A-9494 Sequence 9494, Ap
41 16 61.5 2583 4 US-09-252-991A-9541 Sequence 9541, Ap
42 16 61.5 3096 4 US-09-270-767-13712 Sequence 13712, A
43 16 61.5 60990 4 US-09-949-016-14080 Sequence 14080, A
44 16 61.5 108169 4 US-09-949-016-12898 Sequence 12898, A
45 16 61.5 108169 4 US-09-949-016-15907 Sequence 15907, A

ALIGNMENTS

RESULT 1

US-09-949-016-4238
; Sequence 4238, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4238
; LENGTH: 8865
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4238

Query Match 67.7%; Score 17.6; DB 4; Length 8865;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAAAGCCGCAACTGACGA 26

Db 7087 GGTGTTCCGAGCCGCACTGACCA 7110

RESULT 2

US-08-826-267-1
; Sequence 1, Application US/08826267
; Patent No. 5994070
; GENERAL INFORMATION:
; APPLICANT: Struill, Michel
; TITLE OF INVENTION: No. 5994070el TRIO Molecules and Uses Related Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,267
; FILING DATE: 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,214
; FILING DATE: 27 MARCH (1996)
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8906 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..8647
; US-08-826-267-1

Query Match 67.7%; Score 17.6; DB 2; Length 8906;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGGGTTCCAGCGCAACTGACGA 26
Db 7128 GGTGTTCCGAGCGCCCACTGACCA 7151

RESULT 3
US-09-949-016-15980
; Sequence 15980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15980
; LENGTH: 239527
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(239527)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980

Query Match 67.7%; Score 17.6; DB 4; Length 239527;
Best Local Similarity 83.3%; Pred. No. 1,7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGGGTTCCAGCGCAACTGACGA 26
```

```
Db 221552 GGTGTTCCGAGCGCCCACTGACCA 221575

RESULT 4
5514590-1
; Patent No. 5514590
; APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.; JAMES, ERIC
; TITLE OF INVENTION: EXPRESSION SYSTEM COMPRISING DNA
; ENCODING THE SIGNAL PEPTIDE OF PROTEASE B FROM STREPTOMYCES
; GRISEUS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,644
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 66,938
; FILING DATE: 25-MAY-1993
; APPLICATION NUMBER: 844,937
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: 221,346
; FILING DATE: 18-JUL-1988
; APPLICATION NUMBER: 795,331
; FILING DATE: 06-NOV-1985
; SEQ ID NO:1:
; LENGTH: 1430
5514590-1

Query Match 66.2%; Score 17.2; DB 6; Length 1430;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGTTCACGCGCAACTGACG 25
Db 478 GGGCTCCAGCGCAAACTGACG 499

RESULT 5
5514590-1
; Patent No. 5514590
; APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.; JAMES, ERIC
; TITLE OF INVENTION: EXPRESSION SYSTEM COMPRISING DNA
; ENCODING THE SIGNAL PEPTIDE OF PROTEASE B FROM STREPTOMYCES
; GRISEUS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,644
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 66,938
; FILING DATE: 25-MAY-1993
; APPLICATION NUMBER: 844,937
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: 221,346
; FILING DATE: 18-JUL-1988
; APPLICATION NUMBER: 795,331
; FILING DATE: 06-NOV-1985
; SEQ ID NO:1:
; LENGTH: 1430
5514590-1

Query Match 66.2%; Score 17.2; DB 6; Length 1430;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGTTCACGCGCAACTGACG 25
Db 478 GGGCTCCAGCGCAAACTGACG 499

RESULT 6
US-08-638-931-68/c
; Sequence 68, Application US/08638931
```


Patent No. 6194145
GENERAL INFORMATION:
APPLICANT: HEIDRICH, Bj rn
APPLICANT: ROBINSON, Peter-Nicholas
APPLICANT: TIECKE, Frank
APPLICANT: ROLFS, Arndt
TITLE OF INVENTION: Genus and species-specific identification of
TITLE OF INVENTION: legionella
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikolaide, Marmelstein, Murray & Oram LLP
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/638,931
FILING DATE: 25-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 195 15 891.1
FILING DATE: 29-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Robert B.
REGISTRATION NUMBER: 22,980
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NOS: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 366 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Legionella wadsworthii
STRAIN: 81-716A
INDIVIDUAL ISOLATE: 49wad
US-08-638-931-68
Query Match 65.4%; Score 17; DB 3; Length 366;
Best Local Similarity 80.0%; Pred. No. 97;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 1 AAGGGTTCCAAAGCGCAACTGACG 25
Db 305 AGGTGTTCCAAAGCGCTATTGTGC 281
RESULT 7
US-09-252-991A-2320
Sequence 2320, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 2320
LENGTH: 1116
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2320
Query Match 65.4%; Score 17; DB 4; Length 1116;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 2 AGGGTTCCAAAGCGCAACTGACGA 26
Db 782 AGGTGTTCCCGCGCCACGACGA 806
RESULT 8
US-09-252-991A-2596/c
Sequence 2596, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2596
LENGTH: 2082
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2596
Query Match 65.4%; Score 17; DB 4; Length 2082;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 2 AGGGTTCCAAAGCGCAACTGACGA 26
Db 1358 AGGTGTTCCCGCGCCACGACGA 1334
RESULT 9
US-09-252-991A-2514/c
Sequence 2514, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2514
LENGTH: 2388
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2514
Query Match 65.4%; Score 17; DB 4; Length 2388;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 2 AGGGTTCCAAAGCGCAACTGACGA 26

Db 1269 AGGGTTCCCGCGCGCAACGACGA 1245
||||| ||||| ||||| ||||| |||||

RESULT 10
US-08-822-445-11
; Sequence 11, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..11208

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

US-08-822-445-11

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/396,540
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/822,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 12225 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 190..11208

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

US-09-396-540-11

Query Match 65.4%; Score 17; DB 3; Length 12225;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 AGGGTTCCAGCGCAACTGACGA 26
||| ||||| ||||| ||||| |||||
Db 7969 AGTGGTTCCAAGCAGCAATGGACTA 7993

RESULT 12
US-08-822-445-9
; Sequence 9, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

US-08-822-445-9

```
ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
; LENGTH: 12616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..11592
US-08-822-445-9

Query Match      65.4%; Score 17; DB 2; Length 12616;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      2 AGGGTTCCAAGCGCAACTGACGA 26
Db      8356 AGTGGTTCCAAGCAGCAATGGACTA 8380

RESULT 13
US-09-396-540-9
; Sequence 9, Application US/09396540
; Patent No. 6310182
GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Petrou, Charles
; APPLICANT: Moore, Karen
TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/396,540
; FILING DATE:
CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/822,445
; FILING DATE:
ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
; LENGTH: 12616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
MOLECULE TYPE: DNA
FEATURE:
; NAME/KEY: CDS
; LOCATION: 190..11592
US-09-396-540-9

Query Match      65.4%; Score 17; DB 3; Length 12616;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      2 AGGGTTCCAAGCGCAACTGACGA 26
Db      8356 AGTGGTTCCAAGCAGCAATGGACTA 8380

RESULT 14
US-09-902-540-1185
; Sequence 1185, Application US/09902540
; Patent No. 6833447
GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1185
LENGTH: 17247
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(17247)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1185

Query Match      63.8%; Score 16.6; DB 4; Length 17247;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AAGGGTTCCAAGCGCAACTGA 23
Db      6707 AAGGACTTCCAAGCGCAACTGA 6729

RESULT 15
US-09-949-016-13727
; Sequence 13727, Application US/09949016
; Patent No. 6812339
GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13727
LENGTH: 49931
TYPE: DNA
ORGANISM: Human
```

```
;
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13729

Query Match      63.8%; Score 16.6; DB 4; Length 49931;
Best Local Similarity 82.6%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGGGTTCCAGCGCAACTGAC 24
   |||||||
Db 23506 AGCACTTCCAGCAGCAACTGAC 23528

RESULT 16
US-09-949-016-13728
; Sequence 13728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13728
; LENGTH: 49931
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13728

Query Match      63.8%; Score 16.6; DB 4; Length 49931;
Best Local Similarity 82.6%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGGGTTCCAGCGCAACTGAC 24
   |||||||
Db 23506 AGCACTTCCAGCAGCAACTGAC 23528

RESULT 17
US-09-949-016-13729
; Sequence 13729, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13729
```

```
;
;
; LENGTH: 49931
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13729

Query Match      63.8%; Score 16.6; DB 4; Length 49931;
Best Local Similarity 82.6%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AGGGTTCCAGCGCAACTGAC 24
   |||||||
Db 23506 AGCACTTCCAGCAGCAACTGAC 23528

RESULT 18
US-09-621-976-1947/C
; Sequence 1947, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1947
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 333..482
; NAME/KEY: sig_peptide
; LOCATION: 333..443
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 3.79999995231628
; OTHER INFORMATION: seq LHSFCLSTDCLS/QR
US-09-621-976-1947

Query Match      63.1%; Score 16.4; DB 4; Length 1013;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGGGTTCCAGCGCGCAA 19
   |||||||
Db 374 AGGGTTCCAGCGCGCAA 357

RESULT 19
US-09-620-312D-352
; Sequence 352, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yongnong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
```

```

; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PE_FL_genes Version 1.0
; SEQ ID NO 352
; LENGTH: 4495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (307)..(3006)
US-09-620-312D-352

```

```

Query Match 63.1%; Score 16.4; DB 4; Length 4495;
Best Local Similarity 76.9%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 AAGGGTTCCAAAGCCGCAACTGACGA 26
|||
Db 184 AAGGGTTCCATGCGCAATGAGGA 209
|||

```

RESULT 20

```

US-09-949-016-16329/c
; Sequence 16329, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16329
; LENGTH: 113538
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(113538)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16329

```

```

Query Match 63.1%; Score 16.4; DB 4; Length 113538;
Best Local Similarity 76.9%; Pred. No. 5.7e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 AAGGGTTCCAAAGCCGCAACTGACGA 26
|||
Db 58203 AAGGGTTGCCAAACAGCAATGCAA 58178
|||

```

RESULT 21

```

US-09-103-840A-2

```

```

; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

```

Query Match 63.1%; Score 16.4; DB 3; Length 4403765;
Best Local Similarity 76.9%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 AAGGGTTCCAAAGCCGCAACTGACGA 26
|||
Db 3477587 AAAGTTTCCAAAGTCCGCAAGTGCGA 3477612
|||

```

RESULT 22

```

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

```

Query Match 63.1%; Score 16.4; DB 3; Length 4403765;
Best Local Similarity 76.9%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 AAGGGTTCCAAAGCCGCAACTGACGA 26
|||
Db 1341682 AAAGTTTCCAAAGTCCGCAAGTGCGA 1341657
|||

```

RESULT 23

```

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:

```

```

; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1

Query Match      63.1%; Score 16.4; DB 3; Length 4411529;
Best Local Similarity 76.9%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 AAGGGGTTCCAAAGCGCGCAACTGACGA 26
Db      3481837 AAAGCTTTCCAAAGTCGCAAGTGTCGA 3481862

RESULT 24
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1

Query Match      63.1%; Score 16.4; DB 3; Length 4411529;
Best Local Similarity 76.9%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 AAGGGGTTCCAAAGCGCGCAACTGACGA 26
Db      1342214 AAAGCTTTCCAAAGTCGCAAGTGTCGA 1342189

RESULT 25
US-09-248-796A-709/c
; Sequence 709, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409

```

```

; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 709
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-709

Query Match      62.3%; Score 16.2; DB 4; Length 897;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AAGGGGTTCCAAAGCGCGCAACT 21
Db      155 AAGGGGTTCCAAAGTGCAACT 135

Search completed: March 11, 2005, 13:12:52
Job time : 30.5316 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 129.89 Seconds
(without alignments)

1190.710 Million cell updates/sec

Title: US-09-674-277-10

Perfect score: 26

Sequence: 1 aaggggttcgaagccgaactgacga 26

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	73.1	257	18	US-10-425-115-120301
2	18	69.2	280	17	US-10-437-963-6558
3	18	69.2	350	17	US-10-424-599-1898
4	18	69.2	746	16	US-10-029-386-20439
5	18	69.2	1456	16	US-10-437-963-20793
6	18	69.2	1924	16	US-10-361-460-19
7	18	69.2	2936	18	US-10-425-115-4223
8	17.8	68.5	8400	15	US-10-017-161-1003
9	17.8	68.5	8400	17	US-10-292-798-853
10	17.6	67.7	215	9	US-09-864-761-29907
11	17.6	67.7	513	9	US-09-864-761-13346

12	17.6	67.7	639	17	US-10-424-599-132352
13	17.6	67.7	3731	9	US-09-925-399-114
14	17.6	67.7	3731	10	US-09-925-299-114
15	17.6	67.7	4354	18	US-10-437-963-49722
16	17.6	67.7	8906	17	US-10-374-979-85
17	17.6	67.7	8906	17	US-10-331-496A-87
18	17.6	67.7	8906	17	US-10-182-936A-85
19	17.6	67.7	8906	18	US-10-477-238A-664
20	17.6	67.7	8906	18	US-10-680-287A-664
21	17.2	66.2	257	17	US-10-424-599-75224
22	17.2	66.2	1191	9	US-09-770-445-47
23	17.2	66.2	2339	18	US-10-723-860-6710
24	17	65.4	230	14	US-10-125-258-69
25	17	65.4	564	18	US-10-437-963-86377
26	17	65.4	570	13	US-10-027-632-321893
27	17	65.4	570	17	US-10-027-632-321893
28	17	65.4	701	18	US-10-425-115-135618
29	17	65.4	879	17	US-10-282-122A-15108
30	17	65.4	1035	18	US-10-425-115-135619
31	17	65.4	1293	17	US-10-369-493-44383
32	17	65.4	3036	18	US-10-437-963-37943
33	17	65.4	12225	9	US-09-927-668-11
34	17	65.4	12616	9	US-09-927-668-9
35	17	65.4	13449	10	US-09-873-367C-260
36	17	65.4	13449	15	US-10-171-581-337
37	17	65.4	13449	17	US-10-439-703-28
38	17	65.4	13449	18	US-10-755-889-776
39	16.8	64.6	1250	13	US-10-027-632-58999
40	16.8	64.6	1250	17	US-10-027-632-58999
41	16.6	63.8	25	19	US-10-719-900-17121
42	16.6	63.8	25	19	US-10-719-900-17122
43	16.6	63.8	201	18	US-10-719-993-49799
44	16.6	63.8	201	18	US-10-719-993-49800
45	16.6	63.8	201	18	US-10-719-993-49853

ALIGNMENTS

RESULT 1

US-10-425-115-120301
; Sequence 120301, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 120301
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_41200C.1
US-10-425-115-120301

Query Match 73.1%; Score 19; DB 18; Length 257;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGGTTCACGCGCACTG 22
|||||
Db 171 GGGTTCACGCGCACTG 189

RESULT 2

US-10-437-963-6558

; Sequence 6558, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 6558
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13237C.1
US-10-437-963-6558

Query Match 69.2%; Score 18; DB 18; Length 280;
Best Local Similarity 80.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCCGCAACTGACGA 26
|||||
Db 165 AAGGAGCTCCAAAGTCCCATCTGACGA 190

RESULT 3
US-10-424-599-1898
; Sequence 1898, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 1898
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101719C.1
US-10-424-599-1898

Query Match 69.2%; Score 18; DB 17; Length 350;
Best Local Similarity 80.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCCGCAACTGACGA 26
|||||
Db 259 AAGGGGTTCCAAAGCCCACTATCAA 284

RESULT 4
US-10-029-386-20439
; Sequence 20439, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20439
; LENGTH: 746
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049779.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EST HUMAN HIT: T84960.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P10162, EVALUE 2.70e-02
; OTHER INFORMATION: NT HIT: g114784246, EVALUE 0.00e+00
US-10-029-386-20439

Query Match 69.2%; Score 18; DB 16; Length 746;
Best Local Similarity 80.8%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCCGCAACTGACGA 26
|||||
Db 668 AAGGGGCTCACAGCAGCAGCTGACGA 693

RESULT 5
US-10-437-963-20793
; Sequence 20793, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 20793
; LENGTH: 1456
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26126C.1
US-10-437-963-20793

Query Match 69.2%; Score 18; DB 18; Length 1456;
Best Local Similarity 80.8%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCCGCAACTGACGA 26
|||||
Db 1190 AAGGAGCTCCAAAGTCCCATCTGACGA 1215

RESULT 6
US-10-361-460-19


```
; Sequence 19, Application US/10361460
; Publication No. US20030163839A1
; GENERAL INFORMATION:
; APPLICANT: Helentjaris, Timothy G.
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Encoding Enzymes for Lignin
; FILE REFERENCE: 0709
; CURRENT APPLICATION NUMBER: US/10/361,460
; PRIOR FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/057,082
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)...(1692)
US-10-361-460-19

Query Match      69.2%; Score 18; DB 16; Length 1924;
Best Local Similarity 80.8%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGTTCCAAAGCGCAACTGACGA 26
    ||||| ||||| ||||| ||||| |||||
Db 1390 AAGGGCTTCCAAAGTCGCCCTCGCGA 1415

RESULT 7
US-10-425-115-4293
; Sequence 4293, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 4293
; LENGTH: 2936
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_103914C.1
US-10-425-115-4293

Query Match      69.2%; Score 18; DB 18; Length 2936;
Best Local Similarity 80.8%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGTTCCAAAGCGCAACTGACGA 26
    ||||| ||||| ||||| ||||| |||||
Db 1792 AAGGGCTTCCAAAGTCGCCCTCGCGA 1817

RESULT 8
US-10-017-161-1003/c
; Sequence 1003, Application US/10017161
; Publication No. US20030143669A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
```

```
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1003
; LENGTH: 8400
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(8400)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(715)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5183)..(5293)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6024)..(6123)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8087)..(8200)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (2442)..(2541)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (8231)..(8330)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-017-161-1003

Query Match      68.5%; Score 17.8; DB 15; Length 8400;
Best Local Similarity 90.5%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGTTCCAAGCGCAACTGAC 24
    ||||| ||||| ||||| ||||| |||||
Db 7903 GGGTTCCAAGCGCACCTGCC 7883

RESULT 9
US-10-292-798-853/c
; Sequence 853, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 853
; LENGTH: 8400
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
```

LOCATION: (1)..(8400)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(715)
FEATURE:
NAME/KEY: CDS
LOCATION: (5183)..(5293)
FEATURE:
NAME/KEY: CDS
LOCATION: (6024)..(6123)
FEATURE:
NAME/KEY: CDS
LOCATION: (8087)..(8200)
FEATURE:
NAME/KEY: modified base
LOCATION: (2442)..(2541)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (8231)..(8330)
OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-853

Query Match 68.5%; Score 17.8; DB 17; Length 8400;
Best Local Similarity 90.5%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGTTCACGCGCAACTGAC 24
|||||
Db 7903 GGGTTCACGCGCACTGCC 7883

RESULT 10

US-09-864-761-29907/c
; Sequence 29907, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29907
; LENGTH: 215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010352.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
; OTHER INFORMATION: EST HUMAN HIT: BF685214.1, EVALUE 1.00e-118
; OTHER INFORMATION: NT HIT: G16005921, EVALUE 1.00e-118
; OTHER INFORMATION: SWISSPROT HIT: O75962, EVALUE 2.00e-36
US-09-864-761-29907

Query Match 67.7%; Score 17.6; DB 9; Length 215;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGGTTCCAGCGCGCACTGACGA 26
|||||
Db 111 GGTGTTCCGAGCGCCACTGACCA 88

RESULT 11

US-09-864-761-13346/c
; Sequence 13346, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13346
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
US-09-864-761-13346

Query Match          67.7%; Score 17.6; DB 9; Length 513;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAGCCGCAACTGACGA 26
Db 204 GGTGTTCCGAGCGCCCACTGACCA 181

RESULT 12
US-10-424-599-132352
; Sequence 132352, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; TITLE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132352
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(639)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: PAT_WRT3847_90520C.1
US-10-424-599-132352

Query Match          67.7%; Score 17.6; DB 17; Length 639;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAGCCGCAACTGACGA 26
Db 392 GGGGTTCCAGGGCGCCCACTGACCA 415

RESULT 13
US-09-925-299-114
; Sequence 114, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 3731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3730)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-114

Query Match          67.7%; Score 17.6; DB 9; Length 3731;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAGCCGCAACTGACGA 26
Db 396 GGTGTTCCGAGCGCCCACTGACCA 419

RESULT 14
US-09-925-299-114
; Sequence 114, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 3731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3730)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-114

Query Match          67.7%; Score 17.6; DB 10; Length 3731;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAGCCGCAACTGACGA 26
Db 396 GGTGTTCCGAGCGCCCACTGACCA 419

RESULT 15
US-10-437-963-49722/c
; Sequence 49722, Application US/10437963
```

```

; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: FRANTZ, GRETCHEN
; APPLICANT: HILLAN, KENNETH J.
; APPLICANT: PHILLIPS, HEIDI S.
; APPLICANT: POLAKIS, PAUL
; APPLICANT: SMITH, VICTORIA
; APPLICANT: SPENCER, SUSAN D.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WU, THOMAS D.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5014R1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: US 60/351,885
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/360,066
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/362,004
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: US 60/366,284
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/368,679
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 95
; SEQ ID NO 87
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-331-496A-87

Query Match 67.7%; Score 17.6; DB 17; Length 8906;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAGCCGCACTGACGA 26
Db 7128 GGTGTTCCGAGCCGCACTGACCA 7151

RESULT 18
US-10-182-936A-85
; Sequence 85, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216

; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 49722
; LENGTH: 4354
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5227C.1
; US-10-437-963-49722

Query Match 67.7%; Score 17.6; DB 18; Length 4354;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAGGGGTTCCAGCCGCACTGAC 24
Db 1311 AAGGTTCTTCGAGCGCACTGAC 1288

RESULT 16
US-10-374-979-85
; Sequence 85, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 85
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-374-979-85

Query Match 67.7%; Score 17.6; DB 17; Length 8906;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGGGTTCCAGCCGCACTGACGA 26
Db 7128 GGTGTTCCGAGCCGCACTGACCA 7151

RESULT 17
US-10-331-496A-87
; Sequence 87, Application US/10331496A
; Publication No. US20030228305A1
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-85

Query Match      67.7%; Score 17.6; DB 17; Length 8906;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      3 GGGGTTCCAAAGCGCCCAACTGACGA 26
Db      7128 GGTGTTCCGAGCGCCCACTGACCA 7151

RESULT 19
US-10-477-238A-664
; Sequence 664, Application US/10477238A
; Publication No. US20040221326A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; CURRENT FILING DATE: 2003-11-10
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 664
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-477-238A-664

Query Match      67.7%; Score 17.6; DB 18; Length 8906;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      3 GGGGTTCCAAAGCGCCCAACTGACGA 26
Db      7128 GGTGTTCCGAGCGCCCACTGACCA 7151

RESULT 20
US-10-680-287A-664
; Sequence 664, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: Babij, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
```

```
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 664
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-680-287A-664

Query Match      67.7%; Score 17.6; DB 18; Length 8906;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      3 GGGGTTCCAAAGCGCCCAACTGACGA 26
Db      7128 GGTGTTCCGAGCGCCCACTGACCA 7151

RESULT 21
US-10-424-599-75224
; Sequence 75224, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 75224
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38940C.1
US-10-424-599-75224

Query Match      66.2%; Score 17.2; DB 17; Length 257;
Best Local Similarity 86.4%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 AAGGGTTCCAAAGCGCCCAACTG 22
Db      16 AAGGGTTGCAAGCACCACCACTG 37

RESULT 22
US-09-770-445-47
; Sequence 47, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorchach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Raines, Jennifer L.
; APPLICANT: Price, Jennifer M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Naja
; APPLICANT: Slader, Ted
```

```
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1191)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-445-47

Query Match      66.2%; Score 17.2; DB 9; Length 1191;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGGGTTCCAAAGCGCAACTGA 23
| | | | | | | | | | | | | | | |
Db 402 AGGGGTTCCAGGTCGCAACTCA 423

RESULT 23
US-10-723-860-6710/c
; Sequence 6710, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; METHODS OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882,0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6710
; LENGTH: 2339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-6710

Query Match      66.2%; Score 17.2; DB 18; Length 2339;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GGGGTTCCAAAGCGCAACTGAC 24
| | | | | | | | | | | | | | | |
Db 1440 GAGGTTCCAAAGTTGCAACTGAC 1419

RESULT 24
US-10-125-258-69/c
; Sequence 69, Application US/10125258
; Publication No. US20030028920A1
; GENERAL INFORMATION:
; APPLICANT: Altier, Daniel J.
; APPLICANT: Herrmann, Rafael
; APPLICANT: Lu, Albert L.
; APPLICANT: McCutchen, Billy F.
```

```
; APPLICANT: Presnail, James K.
; APPLICANT: Weaver, Janine L.
; APPLICANT: Wong, James F. H.
; TITLE OF INVENTION: Antimicrobial Polypeptides and Their
; USES
; FILE REFERENCE: 35718/246215
; CURRENT APPLICATION NUMBER: US/10/125,258
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/285,355
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 230
; TYPE: DNA
; ORGANISM: Manduca sexta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(135)
US-10-125-258-69

Query Match      65.4%; Score 17; DB 14; Length 230;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCGCAACTGACG 25
| | | | | | | | | | | | | | | |
Db 29 AAGGGATTCCAGGTGCGCTGAAG 5

RESULT 25
US-10-437-963-86377
; Sequence 86377, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 86377
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(564)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85423C.1
US-10-437-963-86377

Query Match      65.4%; Score 17; DB 18; Length 564;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGGGGTTCCAAAGCGCAACTGACGA 26
| | | | | | | | | | | | | | | |
Db 329 AAGTGGTCCAAAGCGGTAAGTGA 353

Search completed: March 12, 2005, 00:25:14
Job time : 130.89 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 344.049 Seconds
(without alignments)
2876.537 Million cell updates/sec

Title: US-09-674-277-10
Perfect score: 26
Sequence: 1 aaggggtccaagcgcaactgacga 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19.2	73.8	955	7 W29474	W29474 mb99f10.r1
C 2	18.8	72.3	337	5 B0216934	B0216934 603758737
C 3	18.8	72.3	821	4 B1872872	B1872872 603397485
C 4	18.8	72.3	824	9 CC905643	CC905643 t024j14ba
C 5	18.8	72.3	1042	9 CNS02F3H	AL194534 Tetraodon
C 6	18.6	71.5	350	6 CB781359	CB781359 AMGNNUC.C
C 7	18.6	71.5	573	4 B1951276	B1951276 HVSME1002
C 8	18.6	71.5	742	6 CD778991	CD778991 EST650352
C 9	18.6	71.5	742	7 CN245040	CN245040 EST010921
C 10	18.6	71.5	900	6 CD794406	CD794406 EST665767
C 11	18.6	71.5	905	6 CD778990	CD778990 EST650351
C 12	18.6	71.5	949	6 CD794405	CD794405 EST665766
C 13	18.6	71.5	1447	2 BE962423	BE962423 601655714
C 14	18.4	70.8	429	1 AL921426	AL921426 AL921426
C 15	18.4	70.8	748	7 CV064706	CV064706 WNEL14b3
C 16	18.4	70.8	810	7 CK146993	CK146993 AGENCOURT
C 17	18.4	70.8	831	7 CK362212	CK362212 AGENCOURT
C 18	18.4	70.8	903	4 B1093748	B1093748 602860456
C 19	18.2	70.0	378	9 CR123873	CR123873 Reverse s
C 20	18.2	70.0	457	8 AZ048913	AZ048913 GSSBRu031
C 21	18.2	70.0	791	9 CL936731	CL936731 OA_ABA004
C 22	18.2	70.0	965	5 BQ925391	BQ925391 AGENCOURT
C 23	18.2	70.0	1038	9 CL035478	CL035478 CH216-19N
C 24	18	69.2	174	1 AA190947	AA190947 zp83f03.r

C	25	18	69.2	244	6	CD125553	CD125553 ME1-0096T
	26	18	69.2	276	2	BF878054	BF878054 CM2-ET012
	27	18	69.2	383	4	BM500665	BM500665 PAC000000
	28	18	69.2	409	9	AG021986	AG021986 Oryza sat
C	29	18	69.2	412	7	CK468268	CK468268 939646 MA
C	30	18	69.2	414	2	AW845594	AW845594 MRO-0T006
C	31	18	69.2	419	6	CD164561	CD164561 ML1-0089T
	32	18	69.2	438	8	AQ150645	AQ150645 HS_3197_B
	33	18	69.2	449	1	AU173532	AU173532 AU173532
C	34	18	69.2	451	7	H44560	H44560 YG75G10.r1
	35	18	69.2	488	5	BP647639	BP647639 BP647639
C	36	18	69.2	502	5	BQ825787	BQ825787 1030130A0
	37	18	69.2	505	5	BQ825786	BQ825786 1030130A0
	38	18	69.2	509	6	CD164593	CD164593 ML1-0089T
	39	18	69.2	528	6	CD164629	CD164629 ML1-0089T
C	40	18	69.2	548	8	BZ899328	BZ899328 CH240.15L
	41	18	69.2	549	8	CC158024	CC158024 I927e04.b
C	42	18	69.2	563	8	CC080368	CC080368 CSU-K33r
C	43	18	69.2	565	2	BE368858	BE368858 601221488
	44	18	69.2	575	2	BP235574	BP235574 602025396
	45	18	69.2	597	9	CC671080	CC671080 OG7FN70TV

ALIGNMENTS

RESULT 1
LOCUS W29474/c
DEFINITION mb99f10.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:337579.5, similar to gb:X74856 M.musculus L28 mRNA for
ribosomal protein L28 (MOUSE);, mRNA sequence.
W29474
W29474.1 GI:1309619

ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

W29474 955 bp mRNA linear EST 11-SEP-1996
mb99f10.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:337579.5, similar to gb:X74856 M.musculus L28 mRNA for
ribosomal protein L28 (MOUSE);, mRNA sequence.
W29474
W29474.1 GI:1309619
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 955)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyllie, F., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:218979

Seq primer: ETPrimer
High quality sequence stop: 427.
Location/Qualifiers
1..955
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:337579"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/notes="Vector: pT73D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGCGCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 73.8%; Score 19.2; DB 7; Length 955;
 Best Local Similarity 87.5%; Pred. No. 3e+02; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 3;
 Qy 1 AAGGGTTCCAAAGCGCAACTGAC 24
 |||||
 Db 627 AAGGGTTCCAAAGCGCAACTGAC 604

RESULT 2
 LOCUS BU216934/c 337 bp mRNA linear EST 25-NOV-2002
 DEFINITION 603758737F1 CSEQCHN04 Gallus gallus CDNA clone CHEST673j9 5', mRNA sequence.
 ACCESSION BU216934
 VERSION BU216934.1 GI:25397621
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 REFERENCE 1 (bases 1 to 337)
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 TITLE A Comprehensive Collection of Chicken CDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES

Location/Qualifiers
 1..337
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST673j9"
 /tissue_type="whole embryo"
 /dev_stage="20-21"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN04"
 /note="Organ: whole embryo; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. CDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 72.3%; Score 18.8; DB 5; Length 337;

Best Local Similarity 90.9%; Pred. No. 4.2e+02; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 2;
 Qy 2 AAGGGTTCCAAAGCGCAACTGAC 23
 |||||
 Db 25 AAGGTGTTCCAAAGCGCAACTGAC 4

RESULT 3

LOCUS BI872872 821 bp mRNA linear EST 11-OCT-2001
 DEFINITION 603397485F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5401076 5', mRNA sequence.
 ACCESSION BI872872
 VERSION BI872872.1 GI:16046547
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 821)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLAM12022 row: j column: 21
 High quality sequence stop: 690.

FEATURES

Location/Qualifiers
 1..821
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5401076"
 /tissue_type="retina"
 /lab_host="NIH MGC 94"
 /clone_lib="NIH MGC 94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-AT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 72.3%; Score 18.8; DB 4; Length 821;
 Best Local Similarity 90.9%; Pred. No. 4.6e+02; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 2;

Qy 1 AAGGGTTCCAAAGCGCAACTG 22
 |||||

Db 735 AAGGGTTTCACGACCACTG 756
 |||||

RESULT 4

LOCUS CC905643/c 824 bp DNA linear GSS 08-AUG-2003
 DEFINITION t024j14ba.rl TAMBET Bos taurus genomic clone t024j14ba, genomic survey sequence.
 ACCESSION CC905643
 VERSION CC905643.1 GI:33524576
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 824)
 AUTHORS Lin, S., Najjar, F.Z., Adelson, D., Gill, C.A. and Roe, B.A.
 TITLE Bovine BAC End Sequences from Library TAMBT
 JOURNAL Unpublished (2003)
 COMMENT Contact: Bruce A. Roe
 Advanced Center for Genome Technology
 University of Oklahoma Department of Chemistry and Biochemistry
 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 Class: BAC ends
 High quality sequence start: 43
 High quality sequence stop: 355.
 FEATURES
 source
 1. .824
 Location/Qualifiers
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="Angus bull T A M U Shoshone Y6 11519666"
 /db_xref="taxon:9913"
 /clone="t024j14ba"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="TAMBT"
 /notes="Vector: pBeloBAC11; Site 1: HindIII; Site 2: HindIII; TAMBT Bovine BAC library (Male) produced by Texas A&M University, Department of Animal Science."
 ORIGIN
 Query Match 72.3%; Score 18.8; DB 9; Length 824;
 Best Local Similarity 90.9%; Pred. No. 4.6e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAGGGGTTCCAAAGCGCAACTG 22
 |||||
 Db 577 AAGGGGTTCCCTGCGCAACTG 556
 RESULT 5
 CENS02F3H/c
 LOCUS
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
 131P15 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL194534.1 GI:7832640
 VERSION
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 REFERENCE 1
 AUTHORS Roest Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quef, F.,
 Saurin, W. and Weissenbach, J.
 TITLE Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 JOURNAL
 MEDLINE 20296633
 PUBMED 10835645
 REFERENCE 2
 AUTHORS Roest Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quef, F.,
 Saurin, W., Bernot, A. and Weissenbach, J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 JOURNAL
 MEDLINE 20359837
 PUBMED 10899143
 REFERENCE 3 (bases 1 to 1042)
 AUTHORS Genoscope.
 TITLE Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 source
 1. .1042
 Location/Qualifiers
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="131P15"
 /clone_lib="G"
 /notes="Genoscope sequence ID : COAG131CH08LP1-end : T7"
 ORIGIN
 Query Match 72.3%; Score 18.8; DB 9; Length 1042;
 Best Local Similarity 83.3%; Pred. No. 4.8e+02;
 Matches 20; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 GGGGTTCCAAAGCGCAACTGACGA 26
 |||||
 Db 411 GGGGTTCCAGGCGCGACGACGA 388
 RESULT 6
 CB781359/c
 LOCUS
 DEFINITION AMGNNUC.CRI-00004-E3-A Colon Rat 1 (10390) Rattus norvegicus cDNA
 clone cri-00004-e3 5', mRNA sequence.
 ACCESSION CB781359
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 350)
 AUTHORS Angen EST Program.
 TITLE Angen Rat EST Program
 JOURNAL Unpublished (2003)
 COMMENT Contact: Dan Fitzpatrick
 Angen, Inc
 One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 0004 row: e column: 3.
 FEATURES
 source
 1. .350
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="cri-00004-e3"
 /tissue_type="intestine, fetal"
 /clone_lib="Colon Rat 1 (10390)"
 /notes="Vector: pMOB; Site 1: SalI; Site 2: NotI; Colon Rat
 1 Fetal rat intestinal cDNA library."
 ORIGIN
 Query Match 71.5%; Score 18.6; DB 6; Length 350;
 Best Local Similarity 84.0%; Pred. No. 5.3e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AAGGGGTTCCAAAGCGCAACTGACG 25
 |||||
 Db 255 AAGAGGTTCCATCAGCAACTGACG 231
 RESULT 7
 BI951276
 LOCUS
 DEFINITION BI951276 573 bp mRNA linear EST 19-OCT-2001


```

Query Match      71.5%; Score 18.6; DB 6; Length 905;
Best Local Similarity 84.0%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCCGCAACTGACG 25
Db 576 AAGGGGTTCCAAAGCCGCAACTGACG 600

RESULT 12
CD794405
LOCUS EST665766 RAB Rhipicephalus appendiculatus cDNA clone RABG75 3'
DEFINITION end, mRNA sequence.
ACCESSION CD794405
VERSION CD794405.1 GI:49550079
KEYWORDS EST.
SOURCE Rhipicephalus appendiculatus
ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Rhipicephalus.
REFERENCE 1 (bases 1 to 949)
AUTHORS Nene,V., Lee,Y., Skilton,R., Mwaura,S., Quackenbush,J., Gardner,M.
and Bishop,R.
TITLE An index of genes transcribed in the salivary glands of
Rhipicephalus appendiculatus
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: EST665767
Contact: Vishvanath Nene
Parasite Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-610-5968
Fax: 301-838-0208
Email: nene@tigr.org
Seq primer: mixed oligo dt.
FEATURES
source
Location/Qualifiers
1..949
/organism="Rhipicephalus appendiculatus"
/mol_type="mRNA"
/db_xref="taxon:34631"
/dev_stage="Adult"
/lab_host="E. coli strain DH10B-TonA"
/clone_lib="RAB"
/notes="Organ: Salivary glands; Vector: pCMVSPORT6.0.ccdB;
Salivary glands were dissected on day four after
initiation of feeding. Total RNA was prepared using acid
guanidium thiocyanate-phenol-chloroform extraction. The
cDNA library was custom prepared by Invitrogen
Corporation. Briefly, first strand cDNA was primed using
oligo(dt) containing a NotI site. Size fractionated double
stranded cDNA was ligated to EcoRV-NotI cleaved vector and
electroporated into E.coli. Library RAB was made from
ticks infected with Theileria parva."

ORIGIN
Query Match      71.5%; Score 18.6; DB 6; Length 949;
Best Local Similarity 84.0%; Pred. No. 5.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAGGGGTTCCAAAGCCGCAACTGACGA 26
Db 764 AAGGGGTTCCAAAGCCGCAACTGACGA 788

RESULT 13
BE962423
LOCUS BE962423 1447 bp mRNA linear EST 14-DEC-2000
DEFINITION 601655714R1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846290 3',
mRNA sequence.
ACCESSION BE962423

```

```

VERSION BE962423.2 GI:11765310
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1447)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Oct 3, 2000 this sequence version replaced gi:10573128.
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW545 row: 1 column: 03
High quality sequence stop: 279.
Location/Qualifiers
1..1447
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3846290"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_65"
/notes="Organ: Colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 Kb. Library constructed by Life
Technologies."

ORIGIN
Query Match      71.5%; Score 18.6; DB 2; Length 1447;
Best Local Similarity 84.0%; Pred. No. 6.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAGGGGTTCCAAAGCCGCAACTGACG 25
Db 952 AAGGGGTTCCAAAGCCGCAACTGACG 976

RESULT 14
AL921426
LOCUS AL921426 429 bp mRNA linear EST 06-JUL-2004
DEFINITION AL921426 PUR-21+22 Danio rerio cDNA clone 095-G01-2, mRNA sequence.
ACCESSION AL921426
VERSION AL921426.1 GI:23186724
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 429)
AUTHORS Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
Wang,W., Wen,Z. and Peng,J.
TITLE 15000 unique zebrafish EST clusters and their future use in
microarray for profiling gene expression patterns during
embryogenesis
JOURNAL Genome Res. 13 (3), 455-466 (2003)
MEDLINE 22505427
PUBMED 12618376
COMMENT Contact: Peng J
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengj@imcb.a-star.edu.sg
Clone requests: info@openbiosystems.com

```

```

Open Biosystems,
6705 Odyssey Drive, Huntsville, AL 35806.
Location/Qualifiers
1. .429
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="095-G01-2"
/tissue_type="whole embryo or fish"
/dev_stages="mixed stages"
/clone_lib="PUR-21+22"

ORIGIN
Query Match 70.8%; Score 18.4; DB 1; Length 429;
Best Local Similarity 95.0%; Pred. No. 6.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 TTCCAAGCCGCAACTGACGA 26
Db 367 TTCCAAGCCGCAACTGACGA 386

RESULT 15
CV064706 748 bp mRNA linear EST 24-AUG-2004
LOCUS WNEL14b3 Wheat EST endosperm library Triticum aestivum cDNA clone
DEFINITION WNEL14b3 5' similar to Oryza sativa, mRNA sequence.
ACCESSION CV064706
VERSION CV064706.1 GI:51527883
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 748)
Ali, S. Holloway, B. and Taylor, W.C.
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUAM14854 row: c column: 12
High quality sequence start: 12
High quality sequence stop: 748.
Location/Qualifiers
1. .748
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Hartog"
/db_xref="taxon:4565"
/clone="WNEL14b3"
/tissue_type="endosperm"
/dev_stages="developing endosperm tissue 6, 8, 10 dpa (days
post_anthesis)"
/lab_host="DH10B (Life Technology)"
/clone_lib="Wheat EST endosperm library"
/notes="vector: ZiploX; Site 1: Sal I; Site 2: Not I; mRNA
was prepared from endosperm tissues of the wheat cultivar
Hartog. cDNA was synthesised from pooled 6, 8, and 10 dpa
endosperm using Not I-oligo(dT)18 primer/adaptor
(Pharmacia Biotech), and then ligated to the Sal I-Not I
site of ZiploX vector (Life Technology) after adding a
Sal I-Xho I adapter (Stratagene). Constructed by Shahjahan
Ali and Bill Taylor."

ORIGIN
Query Match 70.8%; Score 18.4; DB 7; Length 748;
Best Local Similarity 95.0%; Pred. No. 7.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAGGGTTCCAAGCCGCAAC 20
Db 623 AAGGGTTCCAAGCCGCAATC 642

RESULT 16
CK146993 810 bp mRNA linear EST 04-DEC-2003
LOCUS IMAGE:7063046 5', mRNA sequence.
DEFINITION AGENCOURT 16916979 NCI CGAP_Zemb2 Danio rerio cDNA clone
ACCESSION CK146993
VERSION CK146993.1 GI:38685618
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 810)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbe-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUAM14854 row: c column: 12
High quality sequence start: 12
High quality sequence stop: 738.
Location/Qualifiers
1. .810
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7063046"
/tissue_type="embryo"
/lab_host="DH10B (TI-resistant)"
/clone_lib="NCI CGAP_Zemb2"
/notes="vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI CGAP Library."

ORIGIN
Query Match 70.8%; Score 18.4; DB 7; Length 810;
Best Local Similarity 95.0%; Pred. No. 7.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 TTCCAAGCCGCAACTGACGA 26
Db 414 TTCCAAGCCGCAACTGACGA 433

RESULT 17
CK362212 831 bp mRNA linear EST 23-DEC-2003
LOCUS AGENCOURT_17082929 NIH_ZGC_4 Danio rerio cDNA clone IMAGE:7090591
DEFINITION 5', mRNA sequence.
ACCESSION CK362212
VERSION CK362212.1 GI:40328147
KEYWORDS EST.

```

```

SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 831)
AUTHORS    NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: John Ngai, Univ of CA, Berkeley
            cDNA Library Preparation: Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM14925 row: o column: 05
            High quality sequence start: 19
            High quality sequence stop: 732.

FEATURES   source
            Location/Qualifiers
            1..831
            /organism="Danio rerio"
            /mol_type="mRNA"
            /db_xref="taxon:7955"
            /clone="IMAGE:7090591"
            /lab_host="DH10B Tona"
            /clone_lib="NIH_ZGC_4"
            /notes="Organ: brain/CNS; Vector: pME18S-FL3; Site_1:
            DraIII; Site_2: DraIII"

ORIGIN
Query Match      70.8%; Score 18.4; DB 7; Length 831;
Best Local Similarity 95.0%; Pred. No. 7.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      7 TTCCAAGCCGCAACTGACGA 26
Db      179 TTCCAAGCCGCAACTGACGA 198

RESULT 18
BI093748
LOCUS      602860456F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5001829 5',
DEFINITION mRNA sequence.
VERSION     BI093748.1 GI:14512078
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 903)
AUTHORS    NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM11035 row: o column: 14
            High quality sequence stop: 661.

FEATURES   source
            Location/Qualifiers

SOURCE      1..903
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5001829"
            /cell_line="MGC36"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_10"
            /notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: Not I;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.5 kb. Library prepared by Life
            Technologies."

ORIGIN
Query Match      70.8%; Score 18.4; DB 4; Length 903;
Best Local Similarity 95.0%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 GGTTCCAAGCCGCAACTGAC 24
Db      822 GGTTCCAAGCCGCAACAGAC 841

RESULT 19
CR123873/c
LOCUS      CR123873
DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP92p17, genomic survey sequence.
ACCESSION   CR123873.1 GI:49871325
VERSION     CR123873
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 378)
AUTHORS    Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J., and Bradley,A.
TITLE      Direct Submission
JOURNAL     Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES   source
            Location/Qualifiers
            1..378
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHPP92p17"
            /clone_lib="MHPP"

ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 378;
Best Local Similarity 87.0%; Pred. No. 8.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 GGGGTTCCACGCGCAACTGACG 25
Db      36 GGGGTTCCACGCGCATCTGCCG 14

RESULT 20
AZ048913/c
LOCUS      AZ048913
DEFINITION GSSBrU0310 Sheared genomic library Brucella melitensis biovar
            Abortus genomic clone B2N21, genomic survey sequence.
ACCESSION   AZ048913
VERSION     AZ048913.1 GI:7272828
KEYWORDS    GSS.
SOURCE      Brucella melitensis biovar Abortus (Brucella abortus)
            Brucella melitensis biovar Abortus
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Brucellaceae; Brucella.
REFERENCE   1 (bases 1 to 457)

```

AUTHORS Sanchez,D.O., Zandomeni,R., Cravero,S., Rosetti,O., Grau,O. and Ugalde,R.
TITLE Gene discovery through genomic sequencing survey of the Brucella abortus genome
JOURNAL Unpublished (1999)
COMMENT Contact: Sanchez D.O.
 Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral San Martin)
 Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos Aires, Argentina
 Tel: (54-11) 4580/7255/7
 Fax: (54-11) 4752-9639
 Email: dsanchez@iib.unsam.edu.ar
 Seg primer: T7
 Class: shotgun.

FEATURES

Location/Qualifiers
 1..457

/organism="Brucella melitensis biovar Abortus"
 /mol_type="genomic DNA"
 /strain="S-2308"
 /db_xref="taxon:235"
 /clone="B2N21"
 /note="Vector: pBluescript SK(-) (STRATAGENE); Genomic DNA was mechanically sheared, blunt ended, and size-fractionated by agarose gel electrophoresis. Fragments between 1.5-3 kb were recovered and ligated to the EcoRV site of the pBluescript SK (-) vector."

ORIGIN

Query Match 70.0%; Score 18.2; DB 8; Length 457;
 Best Local Similarity 87.0%; Pred. No. 8.5e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GGGTTCACGCGCAACTGACGA 26
 |||||
 Db 327 GGGTTCACGCGCAACTGCGGA 305

RESULT 21
CL936731
LOCUS CL936731 791 bp DNA linear GSS 14-SEP-2004
DEFINITION OA_ABA0047118.f OA_ABA Oryza australiensis genomic clone
 OA_ABA0047118 5', Genomic survey sequence.

ACCESSION CL936731
VERSION CL936731.1 GI:52067562
SOURCE GSS.

ORGANISM Oryza australiensis
 Oryza australiensis
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (bases 1 to 791)
 Klm,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.

TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu

PCR Primers
 FORWARD: TAA TCA GAC TCA CTA TAG GG
 BACKWARD: CAC TCA TTA GGC ACC CCA
 Plate: 0047 row: I column: 18
 Seg primer: TAA TAC GAC TCA CTA TAG GG
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..791
 /organism="Oryza australiensis"

/mol_type="genomic DNA"
 /db_xref="taxon:4532"
 /clone="OA_ABA0047118"
 /tissue_type="young leaves"
 /lab_host="DH10B T1 phage resistant"
 /clone_lib="OA_ABA"
 /note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 70.0%; Score 18.2; DB 9; Length 791;
 Best Local Similarity 87.0%; Pred. No. 9e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GGGTTCACGCGCAACTGACG 25
 |||||
 Db 370 GGGTTCACGCGCAACTTCCG 392

RESULT 22
BQ925391
LOCUS BQ925391 965 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT 8763855 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6327841 5', mRNA sequence.

ACCESSION BQ925391
VERSION BQ925391.1 GI:22340422
KEYWORDS EST.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 965)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.

CDNA Library Preparation: ResGen, Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LLAM13778 row: b column: 02
 High quality sequence start: 11
 High quality sequence stop: 592.

FEATURES Location/Qualifiers
 1..965
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6327841"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_130"

/note="Organ: otcysts; Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

FEATURES

Location/Qualifiers
 1..965
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6327841"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_130"

/note="Organ: otcysts; Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 70.0%; Score 18.2; DB 5; Length 965;
 Best Local Similarity 87.0%; Pred. No. 9.2e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAGGGTTCACGCGCAACTGA 23
 |||||
 Db 856 AGGGGTTCACGCGCAACTGA 878

RESULT 23
CL035478/c
LOCUS CL035478 1038 bp DNA linear GSS 31-DEC-2003

DEFINITION CH216-39N15.RM1.1 CH216 Xenopus tropicalis genomic clone
CH216-39N15_ genomic survey sequence.
ACCESSION CL035478
VERSION CL035478.1 GI:40488526
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1038)
AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM1 TACGACTCATATAGGAGA
Class: BAC ends
High quality sequence start: 53
High quality sequence stop: 844.
Location/Qualifiers
1..1038
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-39N15"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

FEATURES
source
Query Match 70.0%; Score 18.2; DB 9; Length 1038;
Best Local Similarity 87.0%; Pred. No. 9.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 1 AAGGGTTCCTCAAGCGCAACTGA 23
Db 979 AAGGGTTCCTCAAGCGCAACTGA 957

RESULT 24
AA190947/c
LOCUS AA190947
DEFINITION zp33f03.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
IMAGE:626813 5', mRNA sequence.
ACCESSION AA190947
VERSION AA190947
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 174)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Stepcos, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WASHU-NCI human EST project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LILNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1860 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 116.
Location/Qualifiers
1..174
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5045511"
/db_xref="taxon:9606"
/clone="IMAGE:626813"
/sex="female"
/dev_stage="HeLa S3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene HeLa cell s3 937216"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. HeLa S3
epithelioid carcinoma cells grown to semi-confluency
without induction. Average insert size: 1.5 kb; Uni-ZAP XR
Vector. -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3"

ORIGIN
Query Match 69.2%; Score 18; DB 1; Length 174;
Best Local Similarity 80.8%; Pred. No. 9.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 1 AAGGGTTCCTCAAGCGCAACTGAGA 26
Db 72 AAGGGTTCCTCAAGCGCAACTGAGA 47

RESULT 25
CD125553/c
LOCUS CD125553
DEFINITION ME1-0096T-L257-F06-U.B ME1-0096 Schistosoma mansoni cDNA clone
ME1-0096T-L257-F06.B, mRNA sequence.
ACCESSION CD125553
VERSION CD125553.1 GI:34663567
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni
Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 244)
AUTHORS Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,
Ojopi, E.P.B., Faquiao, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,
Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,
Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,
Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,
Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,
Sa, R.C., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,
Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,
Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.
TITLE Transcription analysis of the acelomate human parasite Schistosoma
mansoni
JOURNAL Nat. Genet. 35 (2), 148-157 (2003)
MEDLINE 22879926
PUBMED 12973350
COMMENT Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquimica
Instituto de Quimica - Universidade de Sao Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjoe@iq.usp.br
This sequence was derived from the PAPESP Schistosoma mansoni EST
Genome Project. All sequences in the project were assembled and
annotated. This entry and all the assembled sequences can be seen
in the following URL: http://bioinfo.iq.usp.br/schisto/
Plate: ME1-0096T-L257 row: 6 column: F.
Location/Qualifiers

source 1. .244
 /organism="Schistosoma mansoni"
 /mol_type="mrna"
 /db_xref="taxon:6183"
 /clone="ME1-0096T-L257-F06.B"
 /sex="mixed pool"
 /dev_stage="egg"
 /lab_host="Mus musculus"
 /clone_lib="ME1-0096"
 /note="Vector: pGEM T-easy"

ORIGIN

Query Match 69.2%; Score 18; DB 6; Length 244;
 Best Local Similarity 80.8%; Pred No. 9.9e+02;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGGGTTCCAGCGCGCAACTGACGA 26
 |||||
 Db 166 AAGGGTTCCAGGACGGAAGCGACGA 141
 |||||

Search completed: March 11, 2005, 13:01:05
 Job time : 347.049 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 38.0362 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-11

Perfect score: 26

Sequence: 1 taagggttcacagcgcaactgacg 26

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04.*
1: Geneseq_1980s.*
2: Geneseq_1990s.*
3: Geneseq_2000s.*
4: Geneseq_2001as.*
5: Geneseq_2001bs.*
6: Geneseq_2002as.*
7: Geneseq_2002bs.*
8: Geneseq_2003as.*
9: Geneseq_2003bs.*
10: Geneseq_2003cs.*
11: Geneseq_2003ds.*
12: Geneseq_2004as.*
13: Geneseq_2004bs.*

Pred. No. is the number of results predicted by chance; to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	26	3 AAZ36111	Aaz36111 Primer de
C 2	26	100.0	1489	3 AAZ36101	Aaz36101 Nucleic a
3	25	96.2	26	3 AAZ36110	Aaz36110 Primer de
C 4	20	76.9	31	3 AAZ36112	Aaz36112 Primer de
C 5	17.8	68.5	6480	10 ABZ67665	Abz67665 Human sec
C 6	17.8	68.5	8400	10 ADC86400	Adc86400 Human GPC
C 7	17.6	67.7	444	12 ADP62800	Adp62800 Maize car
C 8	17.6	67.7	4324	4 ABL16366	Abli16366 Drosophil
C 9	17.2	66.2	215	4 AAK23878	Aak23878 Human bra
C 10	17.2	66.2	215	6 ABS23445	Abbs23445 Human gen
C 11	17.2	66.2	513	4 AAK11215	Aak11215 Human bra
C 12	17.2	66.2	513	6 ABS11018	Abbs11018 Human gen
C 13	17.2	66.2	1191	6 ABN98279	Abn98279 Arabidops
C 14	17.2	66.2	1630	6 AAI99766	Aai99766 Human rib
C 15	17.2	66.2	2321	1 AAN91183	Aan91183 SprA gene
C 16	17.2	66.2	2339	12 ADQ23890	Adq23890 Human sof
C 17	17.2	66.2	3731	3 AAC98104	Aac98104 Human col
C 18	17.2	66.2	4066	12 ADQ67377	Adq67377 Novel hum
C 19	17.2	66.2	4266	12 ADM67130	Adm67130 Human hom
C 20	17.2	66.2	8586	8 ACA57512	Aca57512 Human adi

21	17.2	66.2	8906	2 AAT85392	Aat85392 Human TRI
22	17.2	66.2	8906	5 AAS81211	Aas81211 DNA encod
23	17.2	66.2	8906	5 ABA82699	Abas82699 GTPase ge
24	17.2	66.2	8906	8 ACC46019	Acc46019 Human GTP
25	17.2	66.2	8906	10 ADB98713	Adb98713 Human GTP
26	17.2	66.2	8906	10 ADE82509	Ade82509 Human DNA
27	17.2	66.2	8906	10 ADD89083	Add89083 Encoding
28	17.2	66.2	8906	10 ADG14305	Adg14305 Human TRI
29	17.2	66.2	8908	5 AAS70499	Aas70499 DNA encod
30	17.2	66.2	9041	5 AAS84947	Aas84947 DNA encod
31	17.2	66.2	9437	5 AAS81212	Aas81212 DNA encod
32	17.2	66.2	9449	5 AAS70498	Aas70498 DNA encod
33	17.2	66.2	9449	8 ACD11399	Acd11399 Human DNA
34	17.2	66.2	9450	5 AAS84943	Aas84943 DNA encod
C 35	17.2	66.2	110000	11 ADM27081_05	Continuation (6 of
C 36	17	65.4	230	8 ABQ84722	Abq84722 Manduca s
C 37	17	65.4	746	12 ACH87244	Ach87244 Human gen
C 38	17	65.4	1924	2 AAX25196	Aax25196 Maize 4-c
C 39	17	65.4	4605	4 ABL23432	Abi23432 Drosophil
C 40	17	65.4	5329	4 ABL28224	Abi28224 Drosophil
C 41	17	65.4	6578	10 ADF81808	Adf81808 Leukaemia
C 42	17	65.4	7315	4 AAK52613	Aak52613 Human pol
C 43	17	65.4	8291	4 AAK51629	Aak51629 Human pol
C 44	17	65.4	25911	4 ABL02604	Abi02604 Drosophil
C 45	17	65.4	25929	4 ABL02602	Abi02602 Drosophil

ALIGNMENTS

RESULT 1

AAZ36111
ID AAZ36111 standard; DNA; 26 BP.

XX AAZ36111;

AC AAZ36111;

DT 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

PN WO9955908-A2.

XX 04-NOV-1999.

PD 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

PT Escherichia coli, particularly serotype O157:H7, used for detecting these
bacteria in food.
XX Claim 5; Page 27; 48pp; French.
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
derived from two sequences. The first (AAZ36101) is 99.9% homologous to
the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of *Shigella flexneri*. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 26 BP; 7 A; 7 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGGTTCCAAAGCGCAACTGACG 26
 |||||
 Db 1 TAAGGGGTTCCAAAGCGCAACTGACG 26

RESULT 2
 AAZ36101/c
 ID AAZ36101 standard; DNA; 1489 BP.

AC AAZ36101;

XX 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic *Escherichia coli*.

XX Enterohemorrhagic *Escherichia coli*; EHEC; katP gene; *E. coli* O157:H7;
 XX IS91; ds.

XX *Escherichia coli*.

XX WO9955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT *Escherichia coli*, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.

PS Claim 1; Fig 1; 48pp; French.

XX The present sequence is specific to enterohemorrhagic *Escherichia coli*
 CC (EHEC). The sequence is 99.9% homologous to the katP gene of *E. coli*
 CC O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
 CC homologous with IS91 of *E. coli* (nucleotides 1-406 of the present
 CC sequence). The present sequence is of plasmid origin. Fragments of the
 CC present sequence are used, as probes and primers, for detection of *E.*
 CC *coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC), in human or
 CC animal samples, foods or the environment. The fragments are also useful
 CC for epidemiological studies

XX SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 3; Length 1489;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGGTTCCAAAGCGCAACTGACG 26
 |||||
 Db 427 TAAGGGGTTCCAAAGCGCAACTGACG 402

RESULT 3
 AAZ36110
 ID AAZ36110 standard; DNA; 26 BP.

XX AAZ36110;

XX 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

XX Enterohemorrhagic *Escherichia coli*; EHEC; katP gene; *E. coli* O157:H7;

XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

XX PCR primer; probe; ss.

XX Synthetic.

XX *Escherichia coli*.

XX WO9955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT *Escherichia coli*, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

XX AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic *Escherichia coli* (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of *E. coli* O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of *E. coli* (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of *Shigella flexneri*. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 26 BP; 8 A; 7 C; 8 G; 3 T; 0 U; 0 Other;

Query Match 96.2%; Score 25; DB 3; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.048;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAGGGGTTCCAAAGCGCAACTGACG 26
 |||||
 Db 1 AAGGGGTTCCAAAGCGCAACTGACG 25

RESULT 4
 AAZ36112/c
 ID AAZ36112 standard; DNA; 31 BP.

XX AAZ36112;

XX 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Suwa M, Aesai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
DR P-PSDB; ADC86401.
XX
XX New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX
XX Claim 1; SEQ ID NO 853; 28pp; English.
XX
XX The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC invention.
XX
XX Sequence 8400 BP; 1949 A; 1993 C; 1981 G; 2277 T; 0 U; 200 Other;
SQ
XX
XX Query Match 68.5%; Score 17.8; DB 10; Length 8400;
XX Best Local Similarity 90.5%; Pred. No. 2.8e+02;
XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 5 GGGTTCACAGCGCACTGAC 25
XX |||||
XX DB 7903 GGGTTCACAGCGCACTGCC 7883
XX
XX RESULT 7
XX ADP62800/c
XX ID ADP62800 standard; cDNA; 444 BP.
XX
XX AC ADP62800;
XX
XX DT 09-SEP-2004 (first entry)
XX
XX DE Maize carbon assimilation pathway enzyme cDNA #2813.
XX
XX KW Carbon assimilation pathway enzyme; gene; ss; maize; corn;
XX ribulose-bisphosphate carboxylase; phosphoglycerate kinase;
XX glyceraldehyde 3-phosphate dehydrogenase; triose phosphate isomerase;
XX aldolase; fructose-1,6-bisphosphate; transketolase;
XX sedoheptulose-1,7-bisphosphatase; D-ribulose-5-phosphate-3-epimerase;
XX ribose-5-phosphate isomerase; ribose-5-phosphate kinase;
XX phosphoenolpyruvate carboxylase; NADP-dependent malate dehydrogenase;
XX aspartate aminotransferase; alanine aminotransferase;
XX NADP-dependent malic enzyme; NAD-dependent malic enzyme;
XX PEP carboxykinase; pyruvate; phosphate dikinase; pyrophosphatase;
XX plant metabolic pathway; plant breeding.
XX
XX OS Zea mays.
XX
XX US2004116682-A1.
XX
XX PD 17-JUN-2004.
XX
XX PF 16-NOV-2001; 2001US-00987899.
XX
XX PR 06-MAR-1998; 98US-0076912P.
XX PR 04-MAR-1999; 99US-00262979.
XX
XX (CHEI/) CHEIKH N.
XX (MILL/) MILLER P W.
XX (OCON/) O'CONNELL K M.
XX (LIUJ/) LIU J.
XX

PI Cheikh N, Miller PW, O'connell KM, Liu J;
XX WPI; 2004-498291/47.
XX
XX New substantially purified nucleic acid molecule encoding a maize or
PT soybean carbon assimilation pathway enzyme, useful for gene mapping, gene
PT identification and analysis, plant breeding, and preparation of
PT constructs.
XX
XX Claim 2; SEQ ID NO 5869; 196pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC that encodes a maize or soybean carbon assimilation pathway enzyme or its
CC fragment. The maize or soybean carbon assimilation pathway enzyme or its
CC fragment is selected from ribulose-bisphosphate carboxylase,
CC phosphoglycerate kinase, glyceraldehyde 3-phosphate dehydrogenase, triose
CC phosphate isomerase, aldolase, fructose-1,6-bisphosphatase,
CC transketolase, sedoheptulose-1,7-bisphosphatase, D-ribulose-5-phosphate-3-
CC -epimerase, ribose-5-phosphate isomerase, ribose-5-phosphate kinase,
CC phosphoenolpyruvate carboxylase, NADP-dependent malate dehydrogenase,
CC aspartate aminotransferase, alanine aminotransferase, NADP-dependent
CC malic enzyme, NAD-dependent malic enzyme, PEP carboxykinase, pyruvate,
CC phosphate dikinase and pyrophosphatase. The invention also relates to a
CC substantially purified antibody or its fragment which is capable of
CC specifically binding to a specific maize or soybean carbon assimilation
CC pathway enzyme or its fragment, a transformed plant having a nucleic acid
CC molecule comprising an exogenous promoter region which functions in a
CC plant cell to cause the production of an mRNA molecule, and a method of
CC determining a level or pattern in a plant cell of a carbon assimilation
CC pathway enzyme in a plant metabolic pathway. The methods and compositions
CC of the invention are useful for gene mapping, gene identification and
CC analysis, plant breeding and preparation of constructs for use in plant
CC gene expression and transgenic plants. This sequence represents cDNA
CC encoding a maize carbon assimilation pathway enzyme of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
XX SQ Sequence 444 BP; 116 A; 110 C; 96 G; 122 T; 0 U; 0 Other;
XX
XX Query Match 67.7%; Score 17.6; DB 12; Length 444;
XX Best Local Similarity 83.3%; Pred. No. 2.4e+02;
XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 TAAGGGGTTCACAGCGCACTGA 24
XX |||||
XX DB 359 TAAGGTGATCCACGACGAATTGA 336
XX
XX RESULT 8
XX ABL16366
XX ID ABL16366 standard; DNA; 4324 BP.
XX
XX AC ABL16366;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 571.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX

PA (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 571; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 4324 BP; 1148 A; 951 C; 1076 G; 1149 T; 0 U; 0 Other;
SQ
Query Match 67.7%; Score 17.6; DB 4; Length 4324;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAGGGGTTCCCAAGCGCAACTGAC 25
Db 1605 AAGGGGTTCCCAATCAGCGAATGAC 1628
RESULT 9
AAK23878/c
ID AAK23878 standard; DNA; 215 BP.
AC AAK23878;
XX
XX 05-NOV-2001 (first entry)
DT
DE Human brain expressed single exon probe SEQ ID NO: 23869.
DE
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US0000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
PT
XX

PS Example 4; SEQ ID NO 23869; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
XX Sequence 215 BP; 42 A; 63 C; 58 G; 52 T; 0 U; 0 Other;
SQ
Query Match 66.2%; Score 17.2; DB 4; Length 215;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 GGGGTTCCCAAGCGCAACTGAC 25
Db 111 GGTGTTCCGAGCGCCCACTGAC 90
RESULT 10
ABS23445/c
ID ABS23445 standard; DNA; 215 BP.
XX
XX ABS23445;
AC
XX 19-AUG-2002 (first entry)
DT
XX
XX Human genome-derived single exon probe ORF from lung SEQ ID NO 23436.
DE
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangiomyomatosis; Karagren syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
PT
XX
XX Claim 4; SEQ ID NO 23436; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC

CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying the exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 215 BP; 42 A; 63 C; 58 G; 52 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 6; Length 215;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GGGGTTCCAGCGCACTGAC 25
Db 111 GGTGTTCCGAGCGCCACTGAC 90

RESULT 11
AAK11215/C
ID AAK11215 standard; DNA; 513 BP.
XX AC AAK11215;
XX XX 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe SEQ ID NO: 11206.
XX XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX KW ss.
XX OS Homo sapiens.
XX XX WO200157275-A2.
XX FN 09-AUG-2001.
XX PD 30-JAN-2001; 2001WO-US00006657.
XX PF 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-483446/52.
XX DR Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX PT Example 4; SEQ ID NO 11206; 650pp + Sequence Listing; English.
XX PS The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX SQ Sequence 513 BP; 108 A; 152 C; 135 G; 118 T; 0 U; 0 Other;
Query Match 66.2%; Score 17.2; DB 4; Length 513;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GGGGTTCCAGCGCACTGAC 25
Db 204 GGTGTTCCGAGCGCCACTGAC 183

RESULT 12
ABS11018/c
ID ABS11018 standard; DNA; 513 BP.
XX AC ABS11018;
XX XX 19-AUG-2002 (first entry)
XX DE Human genome-derived single exon probe from lung SEQ ID No 11009.
XX DE Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX KW chronic obstructive pulmonary disease; interstitial lung disease;
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX KW primary ciliary dyskinesia; pulmonary hypertension;
XX KW hyaline membrane disease.
XX OS Homo sapiens.
XX XX WO200186003-A2.
XX XX 15-NOV-2001.
XX XX 30-JAN-2001; 2001WO-US0000665.
XX XX 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX XX (MOLE-) MOLECULAR DYNAMICS INC.
XX FA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2002-114183/15.
XX XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX Claim 1; SEQ ID NO 11009; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridization of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridization to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagazer syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 513 BP; 108 A; 152 C; 135 G; 118 T; 0 U; 0 Other;
Query Match 66.2%; Score 17.2; DB 6; Length 513;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 4 GGGTTCCAGCCCACTGAC 25
Db 204 GGTGTTCCGAGCCCACTGAC 183

XX RESULT 13
XX ABN98279
XX ID ABN98279 standard; DNA; 1191 BP.
XX
XX AC ABN98279;
XX
XX DT 01-AUG-2002 (first entry)
XX
XX DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 47.
XX
XX KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
XX disease; crop; thale cress; tolerance factor; insect; pathogen;
XX nutrition; ds.
XX
XX OS Arabidopsis thaliana.
XX
XX PN US2002023281-A1.

XX 21-FEB-2002.
XX
XX 26-JAN-2001; 2001US-00770445.
XX
XX 27-JAN-2000; 2000US-0178472P.
XX
XX (GORLACH J. GORLACH J.
XX (ANYI) AN Y.
XX (HAMILTON C M. HAMILTON C M.
XX (PRICE J L. PRICE J L.
XX (RAINES T M. RAINES T M.
XX (YU Y. YU Y.
XX (RAMEAKA J G. RAMEAKA J G.
XX (PAGE A. PAGE A.
XX (MATHW A V. MATHW A V.
XX (LEDFO B L. LEDFO B L.
XX (WOESSNER J P. WOESSNER J P.
XX (HAAS W D. HAAS W D.
XX (GARCIA C A. GARCIA C A.
XX (KRICKER M. KRICKER M.
XX (SLATER T. SLATER T.
XX (DAVIS K R. DAVIS K R.
XX (ALLEN K. ALLEN K.
XX (HOFFMAN N. HOFFMAN N.
XX (HURBAN P. HURBAN P.
XX
XX GORLACH J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
XX Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
XX Garcia CA, Kricker M, Slater T, Davis K, Allen K, Hoffman N;
XX Hurban P;
XX WPI; 2002-403163/43.
XX
XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,
XX producing compositions that modulate the expression or function of its
XX encoded protein, and mapping functional regions of protein.
XX
XX Claim 1; SEQ ID NO 47; 49pp + Sequence Listing; English.
XX
XX The invention relates to an Arabidopsis thaliana nucleic acid (I)
XX comprising a sequence capable of hybridising under stringent conditions
XX to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
XX given in the specification or its fragment. A polypeptide (II) encoded by
XX (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
XX genetically modified cell (IV) comprising an exogenous nucleic acid, is
XX useful for screening a candidate agent for its biological effect. (I) is
XX useful in identifying homologous or related genes, in producing
XX compositions that modulate the expression or function of its encoded
XX protein, mapping functional regions of the protein and in studying
XX associated physiological pathways. (I) is also useful for the genetic
XX manipulation of cells, particularly plant cells. (I) is also useful in
XX screening assays of various plant strains to determine the strains that
XX are best capable of withstanding a particular disease or environmental
XX stress. (II) and (III) are useful for screening of biologically active
XX agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
XX pathways. The screened agents are useful in improved methods of treating
XX crops to prevent or treat disease. (II) are also useful in screening
XX programs to identify agents that mimic or enhance the action of tolerance
XX factors. Such agents are useful in improved methods of treating crops to
XX enhance their tolerance to environmental stress. (I) is also useful for
XX enhancing or inhibiting production of a biosynthetic product in a plant.
XX (III) is useful for identifying other mediators that may induce
XX expression of proteins of interest, for establishing the extent to which
XX any specific insect and/or pathogen is responsible for damage to a
XX particular plant, for identifying other mediators that enhance or induce
XX tolerance to environmental stress, for identifying factors involved in
XX biosynthetic pathways of nutritional, commercial, or medicinal value and
XX for identifying productions of nutritional, commercial or medicinal
XX value. (IV) is useful in the study of genetic function and regulation,
XX for alteration of the cellular metabolism and for screening compounds
XX that may affect the biological function of the gene or gene products.
XX Note: The sequence data for this patent did not form part of the printed

```
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=999909770445
XX
SQ Sequence 1191 BP; 272 A; 288 C; 297 G; 314 T; 0 U; 20 Other;
  Query Match 66.2%; Score 17.2; DB 6; Length 1191;
  Best Local Similarity 86.4%; Pred. No. 4.2e+02;
  Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 AGGGTTCCCAAGCGCAACTGA 24
Db 402 AGGGTTCCAGGCTGCAACTCA 423
RESULT 14
AAI99766
ID AAI99766 standard; cDNA; 1630 BP.
XX
AC AAI99766;
XX
DT 22-JAN-2002 (first entry)
XX
DE Human ribosomal S4 protein 12 encoding cDNA.
XX
KW Human; ribosomal S4 protein 12; cytostatic; virucidal; immunomodulatory;
KW antiinflammatory; haemostatic; malignant tumour; HIV; infection;
KW human immunodeficiency virus; immunological disease; gene therapy; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 927..1259
FT /*tag= a
FT /product= "ribosomal S4 protein 12"
FT /note= "claimed in claim 6"
XX
PN WO200175057-A2.
XX
PD 11-OCT-2001.
XX
PP 26-MAR-2001; 2001WO-CN000524.
XX
PR 29-MAR-2000; 2000CN-00115295.
XX
PA (BIOW-) BROWNDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
WPI; 2002-010782/01.
DR P-PSDB; AAMS2107.
XX
PT New polypeptide applicable in diagnosis and treatment of malignant tumor,
PT hemopathy, HIV infection, immunological diseases and inflammation,
PT comprises the human ribosomal S4 protein 12.
XX
PS Claim 6; Page 30-31; 33pp; Chinese.
XX
CC The invention relates to human ribosomal S4 protein 12 with cytostatic,
CC virucidal, immunomodulatory, antiinflammatory and haemostatic activity.
CC The protein and encoding polynucleotide are used in diagnosis and
CC treatment of malignant tumour, haemopathy, human immunodeficiency virus
CC (HIV) infection, immunological diseases and various inflammations. The
CC polynucleotide is useful in gene therapy
XX
SQ Sequence 1630 BP; 404 A; 342 C; 431 G; 453 T; 0 U; 0 Other;
  Query Match 66.2%; Score 17.2; DB 6; Length 1630;
  Best Local Similarity 86.4%; Pred. No. 4.4e+02;
  Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 GGGTTCCCAAGCGCAACTGACG 26
Db 667 GGGTTGCAAGCAGCAACTGAAG 688
```

```
RESULT 15
AAN91183
ID AAN91183 standard; DNA; 2321 BP.
XX
AC AAN91183;
XX
DT 25-MAR-2003 (revised)
DT 07-JUN-1990 (first entry)
XX
DE SprA gene encoding protease A.
XX
KW Protease A; signal peptide; sprA; fusion protein; ss.
XX
OS Streptomyces griseus.
XX
PH Key Location/Qualifiers
FT RBS 176..180
FT /*tag= d
FT sig_peptide 185..298
FT /*tag= a
FT precursor_RNA 299..532
FT /*tag= b
FT mat_peptide 533..1075
FT /*tag= c
FT /product= "protease A"
XX
PN EF300466-A.
XX
PD 25-JAN-1989.
XX
PF 20-JUL-1988; 88EP-00111713.
XX
PR 21-JUL-1987; 87CA-00542678.
XX
PA (CANG-) CANGENE CORP.
XX
PI Garvin RT, Henderson G, Krygsman P, Liu CJ, Davey C, Malek LT;
XX
WPI; 1989-025650/04.
DR P-PSDB; AAP94147.
XX
PT Genes for protease A and protease B from Streptomyces griseus - used for
PT expressing fusion proteins in which protein is expressed in bioactive
PT form.
XX
PS Claim 51; Fig 4; 26pp; English.
XX
CC The sequence is that of a recombinant DNA mol. comprising a signal
CC sequence, and a gene sequence, together encoding a fusion protein of
CC signal peptide-propeptide-protease A. See also AAN91184. (Updated on 25-
CC MAR-2003 to correct DR field.)
XX
SQ Sequence 2321 BP; 370 A; 908 C; 725 G; 318 T; 0 U; 0 Other;
  Query Match 66.2%; Score 17.2; DB 1; Length 2321;
  Best Local Similarity 86.4%; Pred. No. 4.6e+02;
  Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 GGGTTCCCAAGCGCAACTGACG 26
Db 478 GGGTTCCCAAGCGCAACTGACG 499
RESULT 16
ADO23890/c
ID ADO23890 standard; DNA; 2339 BP.
XX
AC ADO23890;
XX
DT 26-AUG-2004 (first entry)
XX
```

DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 6710.
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW db.
XX Homo sapiens.
XX WO2004048938-A2.
XX PN 10-JUN-2004.
XX PD 26-NOV-2003; 2003WO-US038193.
XX PF 26-NOV-2002; 2002US-0429739P.
XX PR (PROT-) PROTEIN DESIGN LABS INC.
XX PA Aziz N, Ginsburg WM, Zlotnik A;
XX PI WPI; 2004-441208/41.
XX DR
XX PT Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX PT and comparing the gene expression, also useful in treating soft tissue
XX PT sarcoma.
XX PT
XX PS Example 2; SEQ ID NO 6710; 210pp; English.
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX CC determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX CC cytostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX CC The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX CC specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 2339 BP; 727 A; 495 C; 656 G; 461 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 12; Length 2339;
Best Local Similarity 86.4%; Pred. No. 4.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAAAGCCGCAACTGAC 25
Db 1440 GAGGTTCCAAAGTTGCAACTGAC 1419

RESULT 17
AAC98104
ID AAC98104 standard; cDNA; 3731 BP.
XX AAC98104;
XX 09-MAR-2001 (first entry)
XX Human colon cancer antigen nucleotide sequence SEQ ID NO:114.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX Homo sapiens.
XX OS

PN WO200055351-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005883.
XX PF 12-MAR-1999; 99US-0124270P.
XX PR (HUMA-) HUMAN GENOME SCI INC.
XX PA Rosen CA, Ruben SM;
XX PI WPI; 2000-587534/55.
XX DR P-PSDB; AAB53347.
XX PT Colon cancer associated gene sequences, referred to as colon cancer
XX PT antigens, useful for the treatment, prevention, and diagnosis of colon
XX PT disorders such as colon cancer.
XX PS Claim 1; Page 552-553; 2104pp; English.
XX CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
XX called human colon cancer antigens, given in AAB53234 to AAB54006. The
XX CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
XX CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
XX CC vulnary, nephrotropic, antiinfective and antibacterial activities, and
XX CC can be used in gene therapy. The colon cancer antigen polynucleotides,
XX CC proteins and antibodies to the proteins are useful for the prevention,
XX CC treatment and diagnosis of colon disorders, such as colon cancer. The
XX CC polynucleotides may be used in diagnostics and research, such as for
XX CC chromosome identification, and as hybridisation probes. The proteins may
XX CC also be used to prevent diseases such as neural disorders, immune system
XX CC disorders, muscular disorders, reproductive disorders, gastrointestinal
XX CC disorders, wounds, renal disorders, infectious diseases, and
XX CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
XX CC sequences used in the exemplification of the present invention
XX SQ Sequence 3731 BP; 938 A; 914 C; 902 G; 971 T; 0 U; 6 Other;

Query Match 66.2%; Score 17.2; DB 3; Length 3731;
Best Local Similarity 86.4%; Pred. No. 4.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAAAGCCGCAACTGAC 25
Db 396 GGTGTTCCGAGCCGCACTGAC 417

RESULT 18
ADQ67377
ID ADQ67377 standard; cDNA; 4066 BP.
XX ADQ67377;
XX 07-OCT-2004 (first entry)
XX Novel human cDNA sequence #2350.
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytoactive; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX Homo sapiens.
XX OS
XX EPI440981-A2.
XX 28-JUL-2004.
XX 21-JAN-2004; 2004EP-00001196.
XX 21-JAN-2003; 2003JP-00102206.
XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

XX P-PSDB; ADQ57684.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 4538; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded

XX polypeptides, sequences hybridizing to these nucleotides, sequences

XX encoding partial polypeptides and sequences having 70% or 90% identity to

XX the nucleotide and protein sequences. The nucleotides and polypeptides

XX are useful as diagnostic markers or therapeutic target for the diseases

XX or morbid states. They are also useful for treating osteoporosis,

XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,

XX dementia and various cancers. This sequence corresponds to a nucleotide

XX sequence of the invention.

SQ Sequence 4066 BP; 992 A; 986 C; 959 G; 1129 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 12; Length 4066;

Best Local Similarity 86.4%; Pred. No. 5e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAGCGCCCACTGAC 25

DB 760 GGTGTTCCGAGCGCCCACTGAC 781

RESULT 19

ADMG67130

ID ADMG67130 standard; DNA; 4266 BP.

XX AC ADMG67130;

XX 03-JUN-2004 (first entry)

DE Human homologue of murine adipocyte specific DNA SeqID 266.

XX human; adipocyte specific; db; adipose tissue; anti-obesity;

XX high mobility group I-C protein; HMGI-C; obesity; leptin; ob; diabetes;

XX adipogenesis; hypertension; cardiovascular disease; anorectic;

XX antidiabetic; hypotensive.

XX Homo sapiens.

OS WO2004011618-A2.

PN 05-FEB-2004.

XX 29-JUL-2003; 2003WO-US023684.

PF 29-JUL-2002; 2002US-0398785P.

XX 12-JUN-2003; 2003US-0478206P.

XX (HMGE-) HMGNE INC.

XX Chada K, Choutinard R, Ashar H, Sayed AMD;

XX WPI; 2004-143846/14.

XX Identifying adipocyte specific genes, useful for treating obesity or

XX diabetes, and for identifying drug targets, by differential gene

XX expression analysis between adipose tissue or stromal vascular tissue of

XX mice of different genotypes.

XX Claim 11; SEQ ID NO 266; 91pp; English.

PS

XX This invention relates to a novel method for identifying genes that are

XX over-expressed in adipose tissue and as such it provides targets for anti

XX -obesity pharmaceutical compositions. Specifically, it refers to a high

XX mobility group I-C protein (HMGI-C) that is associated with obesity and

XX is epistatic to leptin, furthermore, it refers to the ob gene where an

XX autosomal recessive trait is linked to obesity and diabetes. The present

XX invention describes performing differential gene expression analysis

XX between the white adipose tissue (WAT) or stromal vascular tissue (SVT)

XX of any two different mice selected from a group consisting of wild-type,

XX HMGI-C -/-, ob/ob, or HMGI-C -/- ob/ob genotype mice. Accordingly, using

XX this method novel nucleotides and the encoded proteins thereof were

XX identified that are adipocyte specific, and as such can be used for

XX preventing adipogenesis, diagnosing and treating diabetes, obesity,

XX hypertension and cardiovascular disease, as well as screening for

XX compounds that can modulate or prevent adipogenesis and treat diabetes or

XX obesity. These compositions exhibit anorectic, antidiabetic and

XX hypotensive activities. This polynucleotide sequence is a human homologue

XX of a murine adipocyte specific DNA sequence of the invention.

SQ Sequence 4266 BP; 906 A; 1170 C; 995 G; 1195 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 12; Length 4266;

Best Local Similarity 86.4%; Pred. No. 5e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAGCGCCCACTGAC 25

DB 2699 GAGGTTCCAGGTTGCAACTGAC 2720

RESULT 20

ACA57512

ID ACA57512 standard; cDNA; 8586 BP.

XX AC ACA57512;

XX 10-JUN-2003 (first entry)

DE Human adipocyte Selected Interacting domain, SID, cDNA #599.

XX Human; ss; gene; prey; adipocyte; SID; selected interacting domain;

XX anorectic; antidiabetic; protein-protein interaction; diabetes;

XX yeast 2-hybrid assay; metabolic disorder; obesity.

XX Homo sapiens.

OS WO200286122-A2.

PN 31-OCT-2002.

XX 14-MAR-2002; 2002WO-EF003768.

PF 14-MAR-2001; 2001US-0275734P.

XX (HYBR-) HYBRIGENICS.

XX Legrain P, Daviet L;

XX WPI; 2003-103412/09.

XX P-PSDB; ABU70968.

XX New complex between two interacting proteins in adipocyte cells, useful

XX for identifying selected interacting domains that modulate protein

XX interactions, or for preventing or treating metabolic disorders such as

XX obesity or diabetes.

XX Claim 7; Page 301-305; 382pp; English.

PS

XX The invention relates to a complex between two interacting proteins in

XX adipocyte cells, given in the specification. The proteins are identified

XX by selecting a bait protein from a known adipocyte marker and then

XX performing a yeast 2-hybrid selection to isolate prey proteins encoded by

This polynucleotide sequence encodes human TRIO (see AAW27227), a novel phosphoprotein suggested to be a central organiser of multiple signalling

XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological actions. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAG64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 5; Length 8906;
Best Local Similarity 86.4%; Pred. No. 5.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAGCGCAACTGAC 25
||| ||||| ||||| ||||| |||||
Db 7128 GGTGTTCCGAGCGCCCACTGAC 7149

RESULT 23
ID ABA82699 standard; DNA; 8906 BP.
XX ABA82699;
XX 25-JAN-2002 (first entry)
XX GTPase gene SEQ ID NO:85.
XX Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
XX sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
XX antisense therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
XX osteomalacia; fibrous dysplasia; ds.
XX Homo sapiens.
XX WO200177327-A1.
XX 18-OCT-2001.
XX 21-JUN-2000; 2000WO-US016951.
XX 05-APR-2000; 2000US-00543771.
XX 05-APR-2000; 2000US-00544398.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX WPI; 2001-657171/75.
XX New high bone mass (HBM) and Zmax1 genes and proteins useful for
XX modulating bone mass for the treatment of e.g. osteoporosis.
XX Claim 79; Page 386-389; 443pp; English.
XX The present invention describes the human Zmax1 gene and the high bone
XX mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
XX genes have osteopathic activities. The genes can be used in gene therapy,
XX antisense therapy and in the production of vaccines. They can be used in

CC the diagnosis and treatment of bone disorders including osteoporosis,
CC Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
CC ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
CC the exemplification of the present invention
XX SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 5; Length 8906;
Best Local Similarity 86.4%; Pred. No. 5.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAGCGCAACTGAC 25
||| ||||| ||||| ||||| |||||
Db 7128 GGTGTTCCGAGCGCCCACTGAC 7149

RESULT 24
ID ACC46019 standard; DNA; 8906 BP.
XX ACC46019;
XX 02-JUN-2003 (first entry)
XX Human GTPase gene TRIO.
XX Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
XX gene therapy; bone density modulation; bone strength; trabecular number;
XX bone size; bone tissue connectivity; bone disease; osteoporosis;
XX osteomalacia; rickets; Paget's disease; neoplasm of the bone; gene; ds.
XX Homo sapiens.
XX WO200292764-A2.
XX 21-NOV-2002.
XX 13-MAY-2002; 2002WO-US014876.
XX 11-MAY-2001; 2001US-0290071P.
XX 17-MAY-2001; 2001US-0291311P.
XX 01-FEB-2002; 2002US-0353058P.
XX 04-MAR-2002; 2002US-0361293P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX (AMHP) WYETH.
XX Babij P, Bex FU, Yaworsky PJ, Bodine PV;
XX WPI; 2003-129278/12.
XX New transgenic animals (e.g. mice), useful as models for studying bone
XX density modulation, developing drugs for treating or preventing bone
XX diseases (e.g. osteoporosis), or diagnosing diseases characterized by
XX reduced bone density.
XX Disclosure; Page 509-512; 603pp; English.
XX The invention relates to novel transgenic animals expressing the high
XX bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
XX comprising an alteration of the gene encoding LRP5 or LRP6, or expressing
XX an LRP5 that is modulated by an altered gene control sequence introduced
XX by homologous or non-homologous recombination. The transgenic animals are
XX for the study of bone density modulation or bone mass modulation. The
XX invention has osteopathic and cytostatic activity. The polynucleotides of
XX the invention may have a use in gene therapy. The transgenic animals and
XX nucleic acids are for the study of bone density modulation, where the
XX bone mass is modulated relative to non-transgenic animals of the same
XX species in more than one parameter selected from bone density, bone
XX strength, trabecular number, bone size, or bone tissue connectivity. The
XX transgenic animals, nucleic acids and methods are useful for identifying
XX molecules involved in bone development, and for developing pharmaceutical
XX compositions, which may be employed for treating or preventing bone

CC diseases, e.g. osteoporosis, osteomalacia, rickets, Paget's disease, or
CC neoplasms of the bone. The transgenic animals and nucleic acids are also
CC useful in methods for diagnosing diseases involved in bone development,
CC or characterised by reduced bone density or mass. The present sequence is
CC used in the exemplification of the invention

XX SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

Query Match 66.2%; Score 17.2; DB 8; Length 8906;
Best Local Similarity 86.4%; Pred. No. 5.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGTTCCAAAGCGGCACTGAC 25

DB 7128 GGTGTTCCGAGCGGCCTGAC 7149

RESULT 25

ADB98713

ID ADB98713 standard; DNA; 8906 BP.

XX AC ADB98713;

XX DT 04-DEC-2003 (first entry)

XX DE Human GTPase DNA.

XX KW Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmax1; LRP6;

XX KW bone mass modulation; osteoporosis; ds.

XX OS Homo sapiens.

XX PN WO200292000-A2.

XX PD 21-NOV-2002.

XX PF 13-MAY-2002; 2002WO-US014877.

XX PR 11-MAY-2001; 2001US-0290071P.

XX PR 17-MAY-2001; 2001US-0291311P.

XX PR 01-FEB-2002; 2002US-0353058P.

XX PR 04-MAR-2002; 2002US-0361293P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PA (AMHP) WYETH.

XX PI Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;

XX DR WPI; 2003-129214/12.

XX DR N-PSDB; ADB98736.

XX PT New nucleic acid comprising a mutation in LRP5 or LRP6, useful for

XX PT diagnosing a HBM-like phenotype in a subject and for preparing a

XX PT composition for modulating bone mass and/or lipid levels in a subject

XX PT suffering from e.g. osteoporosis.

XX PS Disclosure; SEQ ID NO 664; 629pp; English.

XX CC The present invention relates to High Bone Mass (HBM), LRP5 (Zmax1) and

XX CC LRP6 mutants, which results in a HBM-like phenotype when expressed in a

XX CC cell. The HBM-like phenotype results in bone mass modulation and/or lipid

XX CC level modulation. The invention is useful for diagnosing a HBM-like

XX CC phenotype in a subject and for preparing a composition for modulating

XX CC bone mass and/or lipid levels in a subject suffering from e.g.

XX CC osteoporosis. The present sequence was used to illustrate the invention.

XX SQ Sequence 8906 BP; 2393 A; 2260 C; 2420 G; 1833 T; 0 U; 0 Other;

QY

4 GGGTTCCAAAGCGGCACTGAC 25

Query Match 66.2%; Score 17.2; DB 10; Length 8906;

Best Local Similarity 86.4%; Pred. No. 5.5e+02;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY

4 GGGTTCCAAAGCGGCACTGAC 25

Query Match

Best Local Similarity

Matches

19; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Db 7128 GGTGTTCCGAGCGGCCTGAC 7149

Search completed: March 11, 2005, 04:19:53
Job time : 39.0362 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 10.5316 Seconds
(Without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-11

Perfect score: 26

Sequence: 1 taagggttccaagcgcaactgacg 26

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405588

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.2	66.2	1430	6	Patent No. 5514590-1
2	17.2	66.2	1430	6	Patent No. 5514590-1
3	17.2	66.2	8865	4	Sequence 4238, Ap
4	17.2	66.2	8906	2	Sequence 1, Appl
5	17.2	66.2	239527	4	Sequence 15980, A
6	17	65.4	366	3	Sequence 68, Appl
7	16.8	64.6	601	4	Sequence 113381, A
8	16.8	64.6	601	4	Sequence 113382, A
9	16.8	64.6	285986	4	Sequence 12287, A
10	16.8	64.6	288031	4	Sequence 14864, A
11	16.6	63.8	12225	2	Sequence 14864, A
12	16.6	63.8	12225	2	Sequence 11, Appl
13	16.6	63.8	12616	2	Sequence 9, Appl
14	16.6	63.8	12616	2	Sequence 9, Appl
15	16.6	63.8	12747	4	Sequence 1185, Ap
16	16.6	63.8	28055	4	Sequence 1841, A
17	16.6	63.8	49931	4	Sequence 13727, A
18	16.6	63.8	49931	4	Sequence 13728, A
19	16.6	63.8	49931	4	Sequence 13729, A
20	16.4	63.1	1013	4	Sequence 1947, Ap
21	16.4	63.1	3096	4	Sequence 13712, A
22	16.4	63.1	304042	4	Sequence 13632, A
23	16.2	62.3	897	4	Sequence 709, App
24	16.2	62.3	42574	4	Sequence 17525, A
25	16	61.5	462	2	Sequence 10, Appl
26	16	61.5	462	3	Sequence 10, Appl
27	16	61.5	601	4	Sequence 61467, A

```

c 28 16 61.5 601 4 US-09-949-016-176383,
29 16 61.5 1116 4 US-09-252-991A-2320
30 16 61.5 1678 4 US-09-673-395A-111
31 16 61.5 2037 4 US-09-902-540-6164
32 16 61.5 2039 4 US-09-902-540-6164
33 16 61.5 2082 4 US-09-252-991A-2596
34 16 61.5 2388 4 US-09-252-991A-2514
35 16 61.5 28366 4 US-09-949-016-16763
36 16 61.5 108169 4 US-09-949-016-12898
37 16 61.5 108169 4 US-09-949-016-15907
38 16 61.5 113538 4 US-09-949-016-16329
39 16 61.5 206433 4 US-09-949-016-13527
40 16 61.5 254778 4 US-09-949-016-12417
41 15.8 60.8 1043 4 US-09-023-655-843
42 15.6 60.0 30 3 US-09-242-690A-25
43 15.6 60.0 30 4 US-09-908-855-25
44 15.6 60.0 293 3 US-09-488-295-14
45 15.6 60.0 305 4 US-09-270-767-2748

```

ALIGNMENTS

```

RESULT 1
5514590-1
;Patent No. 5514590
; APPLICANT: GARVIN, ROBERT T.;MALEK, LAWRENCE T.;JAMES, ERIC
; TITLE OF INVENTION: EXPRESSION SYSTEM COMPRISING DNA
;ENCODING THE SIGNAL PEPTIDE OF PROTEASE B FROM STREPTOMYCES
;GRISEUS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,644
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 66,938
; FILING DATE: 25-MAY-1993
; APPLICATION NUMBER: 844,937
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: 221,346
; FILING DATE: 18-JUL-1988
; APPLICATION NUMBER: 795,331
; FILING DATE: 06-NOV-1985
;SEQ ID NO:1:
; LENGTH: 1430
5514590-1

```

```

Query Match 66.2%; Score 17.2; DB 6; Length 1430;
Best Local Similarity 86.4%; Pred.No.96;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 5 GGGTTCGAAGCGCAACTGACG 26
||| ||||| ||| |||||
Db 478 GGGTTCGAAGCGCAACTGACG 499

```

```

RESULT 2
5514590-1
;Patent No. 5514590
; APPLICANT: GARVIN, ROBERT T.;MALEK, LAWRENCE T.;JAMES, ERIC
; TITLE OF INVENTION: EXPRESSION SYSTEM COMPRISING DNA
;ENCODING THE SIGNAL PEPTIDE OF PROTEASE B FROM STREPTOMYCES
;GRISEUS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,644
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 66,938
; FILING DATE: 25-MAY-1993
; APPLICATION NUMBER: 844,937
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: 221,346

```

```

; FILING DATE: 18-JUL-1988
; APPLICATION NUMBER: 795,331
; FILING DATE: 06-NOV-1985
; SEQ ID NO:1:
; LENGTH: 1430
5514590-1

Query Match      66.2%; Score 17.2; DB 6; Length 1430;
Best Local Similarity 86.4%; Pred. No. 96;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 GGGTTCCAAAGCGCAACTGACG 26
   |||||
DB 478 GGGCTCCAAGCGCAAACTGACG 499

RESULT 3
US-09-949-016-4238
; Sequence 4238, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4238
; LENGTH: 8865
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4238

Query Match      66.2%; Score 17.2; DB 4; Length 8865;
Best Local Similarity 86.4%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGTTCCAAAGCGCAACTGAC 25
   |||||
DB 7087 GGTGTCCGAGCGCCACTGAC 7108

RESULT 4
US-08-826-267-1
; Sequence 1, Application US/08826267
; Patent No. 5994070
; GENERAL INFORMATION:
; APPLICANT: Streuli, Michel
; TITLE OF INVENTION: No. 5994070el TRIO Molecules and Uses Related Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/826,267
; FILING DATE: 1997

```

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113381
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-113381

Query Match          64.6%; Score 16.8; DB 4; Length 601;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TAAGGGGTTCCAAGCCGCAA 20
      ||| ||||| ||||| |||||
Db     224 TAAAGGGTTCCAAGCAGCAA 243

RESULT 8
US-09-949-016-113382
; Sequence 113382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113382
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-113382

Query Match          64.6%; Score 16.8; DB 4; Length 601;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TAAGGGGTTCCAAGCCGCAA 20
      ||| ||||| ||||| |||||
Db     365 TAAAGGGTTCCAAGCAGCAA 384

RESULT 9
US-09-949-016-12287/c
; Sequence 12287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12287
; LENGTH: 285986

```

```
;
;
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(285986)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12287

Query Match      64.6%; Score 16.8; DB 4; Length 285986;
Best Local Similarity 90.0%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TAAGGGTTCCAAGCCGCAA 20
Db 207830 TAAAGGGTTCCAAGCAGCAA 207811

RESULT 10
US-09-949-016-14864/c
; Sequence 14864, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14864
; LENGTH: 288031
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(288031)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14864

Query Match      64.6%; Score 16.8; DB 4; Length 288031;
Best Local Similarity 90.0%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TAAGGGTTCCAAGCCGCAA 20
Db 257875 TAAAGGGTTCCAAGCAGCAA 257856

RESULT 11
US-08-822-445-11
; Sequence 11, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,445
FILING DATE: 21-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 12225 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 190..11208
US-08-822-445-11

Query Match      63.8%; Score 16.6; DB 2; Length 12225;
Best Local Similarity 82.6%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGGGTTCCAAAGCGCAACTGAC 25
Db 7969 AGTGGTTCCAAGCAGCAATGAC 7991

RESULT 12
US-09-396-540-11
; Sequence 11, Application US/09396540
; Patent No. 6310182
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/396,540
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 08/822,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
```

TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 12225 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 190..11208
US-09-396-540-11

Query Match 63.8%; Score 16.6; DB 3; Length 12225;
Best Local Similarity 82.6%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGGGTTCCAAGCCGCAACTGAC 25
Db 7969 AGTGGTTCCAAGCAGCAATGGAC 7991
|||||

RESULT 13

US-08-822-445-9
Sequence 9, Application US/08822445
Patent No. 5952223
GENERAL INFORMATION:
APPLICANT: Kaplan, Jerry
APPLICANT: Perou, Charles
APPLICANT: Moore, Karen
TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: 21-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 12616 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 190..11592
US-08-822-445-9

Query Match 63.8%; Score 16.6; DB 2; Length 12616;
Best Local Similarity 82.6%; Pred. No. 2.9e+02;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 3 AGGGTTCCAAGCCGCAACTGAC 25
Db 8356 AGTGGTTCCAAGCAGCAATGGAC 8378
|||||

RESULT 14

US-09-396-540-9
Sequence 9, Application US/09396540
Patent No. 6310182
GENERAL INFORMATION:
APPLICANT: Kaplan, Jerry
APPLICANT: Perou, Charles
APPLICANT: Moore, Karen
TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: US/09396,540
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/822,445
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-062-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 12616 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 190..11592
US-09-396-540-9

Query Match 63.8%; Score 16.6; DB 3; Length 12616;
Best Local Similarity 82.6%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGGGTTCCAAGCCGCAACTGAC 25
Db 8356 AGTGGTTCCAAGCAGCAATGGAC 8378
|||||

RESULT 15

US-09-902-540-1185
Sequence 1185, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.

```
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1185
; LENGTH: 17247
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(17247)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1185

Query Match      63.8%; Score 16.6; DB 4; Length 17247;
Best Local Similarity 82.6%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AAGGGTTCCAAAGCGCAACTGA 24
||||| ||||||| |||||
Db 6707 AAGGACTTCCAAAGCGCAACCTGA 6729

RESULT 16
US-09-949-016-16441/c
; Sequence 16441, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16441
; LENGTH: 28055
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16441

Query Match      63.8%; Score 16.6; DB 4; Length 28055;
Best Local Similarity 82.6%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAAGGGTTCCAAAGCGCAACTG 23
||||| ||||||| |||||
Db 25609 TAAGGGCTTCCAGGTACATCTG 25587

RESULT 17
US-09-949-016-13727
; Sequence 13727, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13727
; LENGTH: 49931
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13727

Query Match      63.8%; Score 16.6; DB 4; Length 49931;
Best Local Similarity 82.6%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGGGTTCCAAAGCGCAACTGAC 25
||||| ||||||| |||||
Db 23506 AGCACTTCCAAAGCAGCAACTGAC 23528

RESULT 18
US-09-949-016-13728
; Sequence 13728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13728
; LENGTH: 49931
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13728

Query Match      63.8%; Score 16.6; DB 4; Length 49931;
Best Local Similarity 82.6%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGGGTTCCAAAGCGCAACTGAC 25
||||| ||||||| |||||
Db 23506 AGCACTTCCAAAGCAGCAACTGAC 23528

RESULT 19
US-09-949-016-13729
; Sequence 13729, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13729
; LENGTH: 49931
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49931)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13729

Query Match 63.1%; Score 16.6; DB 4; Length 49931;
Best Local Similarity 82.6%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGGGGTTCCAGCGCAACTGAC 25
Db 23506 AGCACTTCCAGCAGCAACTGAC 23528

RESULT 20
US-09-621-976-1947/c
; Sequence 1947, Application US/09621976
; Patent No. 6839083
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1947
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 333..482
; NAME/KEY: sig_peptide
; LOCATION: 333..443
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 3.79999995231628
; OTHER INFORMATION: seq LHSFCLSTDCLS/QR
US-09-621-976-1947

Query Match 63.1%; Score 16.4; DB 4; Length 1013;
Best Local Similarity 94.4%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AGGGGTTCCAGCGCAAA 20
Db 374 AGGGGTTCCAGCGCAAA 357

RESULT 21
US-09-270-767-13712
; Sequence 13712, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13712
; LENGTH: 3096
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13712

Query Match 63.1%; Score 16.4; DB 4; Length 3096;
Best Local Similarity 76.9%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TAAAGGGTTCCAAAGCGCAACTGACG 26
Db 1394 TCAGAGAGTTGCCAGCGCAACTGGCG 1419

RESULT 22
US-09-949-016-13632/c
; Sequence 13632, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13632
; LENGTH: 300402
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300402)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13632

Query Match 63.1%; Score 16.4; DB 4; Length 300402;
Best Local Similarity 76.9%; Pred. No. 6.3e+02;
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TAAAGGGTTCCAAAGCGCAACTGACG 26
Db 118429 TAATGGGTTCCATGCGCAGCAACTGAAG 118404

RESULT 23
US-09-248-796A-709/c
; Sequence 709, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409

```
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 709
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-709

Query Match          62.3%; Score 16.2; DB 4; Length 897;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 AAGGGTTCCAAAGCGCAACT 22
Db 155 AAGGGTTCCAAAGTGCAACT 135

RESULT 24
US-09-949-016-17525
; Sequence 17525, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17525
; LENGTH: 42574
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17525

Query Match          62.3%; Score 16.2; DB 4; Length 42574;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 AAGGGTTCCAAAGCGCAACT 22
Db 6907 AAGAGGTTCCAAAGGCAACT 6927

RESULT 25
US-08-743-637B-10/c
; Sequence 10, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: QUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella pneumoniae
; US-08-743-637B-10

Query Match          61.5%; Score 16; DB 2; Length 462;
Best Local Similarity 79.2%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 2 AAGGGTTCCAAAGCGCAACTGAC 25
Db 63 AGGCGGATCAAGCGGCAACTGAC 40

Search completed: March 11, 2005, 13:12:55
Job time : 13:5316 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 129.89 Seconds
(without alignment)
1190.710 Million cell updates/sec

Title: US-09-674-277-11

Perfect score: 26

Sequence: 1 taagggttccaagcgcaactgacg 26

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	73.1	257	18	US-10-425-115-120301
2	18.2	70.0	257	17	US-10-424-599-75224, A
3	17.8	68.5	350	17	US-10-424-599-75224, A
4	17.8	68.5	8400	15	US-10-017-161-1003
5	17.8	68.5	8400	17	US-10-292-798-853
6	17.6	67.7	444	11	US-09-987-899-5869
7	17.6	67.7	1828	13	US-10-027-632-97219
8	17.6	67.7	1828	13	US-10-027-632-97220, A
9	17.6	67.7	1828	17	US-10-027-632-97219
10	17.6	67.7	1828	17	US-10-027-632-97220, A
11	17.6	67.7	4354	18	US-10-437-963-49722

c 12	17.2	66.2	215	9	US-09-864-761-29907	Sequence 29907, A
c 13	17.2	66.2	513	9	US-09-864-761-13346	Sequence 13346, A
c 14	17.2	66.2	639	17	US-10-424-599-132352	Sequence 132352, A
c 15	17.2	66.2	1191	9	US-09-770-445-47	Sequence 47, Appl
c 16	17.2	66.2	2339	18	US-10-723-860-6710	Sequence 6710, Ap
c 17	17.2	66.2	2339	9	US-09-925-299-114	Sequence 114, App
c 18	17.2	66.2	3731	10	US-09-925-299-114	Sequence 114, App
c 19	17.2	66.2	8906	17	US-10-374-979-85	Sequence 85, Appl
c 20	17.2	66.2	8906	17	US-10-381-496A-87	Sequence 87, Appl
c 21	17.2	66.2	8906	17	US-10-182-936A-85	Sequence 85, Appl
c 22	17.2	66.2	8906	18	US-10-477-238A-664	Sequence 664, App
c 23	17.2	66.2	8906	18	US-10-680-287A-664	Sequence 664, App
c 24	17	65.4	230	14	US-10-125-258-69	Sequence 69, Appl
c 25	17	65.4	280	18	US-10-437-963-6558	Sequence 6558, Ap
c 26	17	65.4	746	16	US-10-029-386-20439	Sequence 20439, A
c 27	17	65.4	777	13	US-10-027-632-153646	Sequence 153646, A
c 28	17	65.4	777	13	US-10-027-632-153647	Sequence 153647, A
c 29	17	65.4	777	17	US-10-027-632-153646	Sequence 153646, A
c 30	17	65.4	777	17	US-10-027-632-153647	Sequence 153647, A
c 31	17	65.4	1145	13	US-10-027-632-117630	Sequence 117630, A
c 32	17	65.4	1145	13	US-10-027-632-117631	Sequence 117631, A
c 33	17	65.4	1145	17	US-10-027-632-117630	Sequence 117630, A
c 34	17	65.4	1145	17	US-10-027-632-117631	Sequence 117631, A
c 35	17	65.4	1250	13	US-10-027-632-58999	Sequence 58999, A
c 36	17	65.4	1250	17	US-10-027-632-58999	Sequence 58999, A
c 37	17	65.4	1456	18	US-10-437-963-20793	Sequence 20793, A
c 38	17	65.4	1924	16	US-10-361-460-19	Sequence 19, Appl
c 39	17	65.4	2936	18	US-10-425-115-4293	Sequence 4293, Ap
c 40	16.6	63.8	25	19	US-10-719-900-17121	Sequence 17121, A
c 41	16.6	63.8	25	19	US-10-719-900-17122	Sequence 17122, A
c 42	16.6	63.8	201	18	US-10-719-993-49799	Sequence 49799, A
c 43	16.6	63.8	201	18	US-10-719-993-49800	Sequence 49800, A
c 44	16.6	63.8	201	18	US-10-719-993-49853	Sequence 49853, A
c 45	16.6	63.8	201	18	US-10-719-993-49875	Sequence 49875, A

ALIGNMENTS

RESULT 1

US-10-425-115-120301
; Sequence 120301, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 120301
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_41200C.1
US-10-425-115-120301

Query Match 73.1%; Score 19; DB 18; Length 257;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 GGGTTCCAAGCGCAACTG 23
Dd 171 GGGTTCCAAGCGCAACTG 189

RESULT 2

US-10-424-599-75224

```
; Sequence 75224, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 75224
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38940C.1
US-10-424-599-75224

Query Match          70.0%; Score 18.2; DB 17; Length 257;
Best Local Similarity 87.0%; Pred. No. 91;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TAAGGGTTCCAAAGCGCACTG 23
    ||||| ||||| |||||
Db 15 TAAGGGTTGCAAGCACCACCTG 37

RESULT 3
US-10-424-599-1898
; Sequence 1898, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 1898
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101719C.1
US-10-424-599-1898

Query Match          68.5%; Score 17.8; DB 17; Length 350;
Best Local Similarity 90.5%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAGGGTTCCAAAGCGCACT 22
    ||||| ||||| |||||
Db 259 AAGGGTTCCAAAGCCCCCACT 279

RESULT 4
US-10-017-161-1003/c
; Sequence 1003, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
```

```

; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97219
; LENGTH: 1828
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97219

Query Match          67.7%   Score 17.6;   DB 13;   Length 1828;
Best Local Similarity 83.3%   Pred. No. 1.9e+02;
Matches 20;   Conservative 0;   Mismatches 4;   Indels 0;   Gaps 0

Qy  1  TAAGGGGTTCCAAGCCGCAACTGA 24
      |||||  |||||  |||||  |||||
Db  1439  TAAGGGGTGTCGAAGCAGCATCTGA 1416

RESULT 8
US-10-027-632-97220/c
; Sequence 97220, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97220
; LENGTH: 1828
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97220

Query Match          67.7%   Score 17.6;   DB 13;   Length 1828;
Best Local Similarity 83.3%   Pred. No. 1.9e+02;
Matches 20;   Conservative 0;   Mismatches 4;   Indels 0;   Gaps 0

Qy  1  TAAGGGGTTCCAAGCCGCAACTGA 24
      |||||  |||||  |||||  |||||

```

Db	1439	TAAGGGTGTCCAGCAGCATCTGA	1416	
RESULT 9				
US-10-027-632-97219/c				
; Sequence 97219, Application US/10027632				
; Publication No. US20030204075A9				
; GENERAL INFORMATION:				
; APPLICANT: Wang, David G.				
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide				
; POLYMORPHISMS IN THE HUMAN GENOME				
; FILE REFERENCE: 10827.129				
; CURRENT APPLICATION NUMBER: US/10/027,632				
; CURRENT FILING DATE: 2002-04-30				
; PRIOR APPLICATION NUMBER: US 60/218,006				
; PRIOR FILING DATE: 2000-07-12				
; PRIOR APPLICATION NUMBER: US 60/198,676				
; PRIOR FILING DATE: 2000-04-20				
; PRIOR APPLICATION NUMBER: US 60/193,483				
; PRIOR FILING DATE: 2000-03-29				
; PRIOR APPLICATION NUMBER: US 60/185,218				
; PRIOR FILING DATE: 2000-02-24				
; PRIOR APPLICATION NUMBER: US 60/167,363				
; PRIOR FILING DATE: 1999-11-23				
; PRIOR APPLICATION NUMBER: US 60/156,358				
; PRIOR FILING DATE: 1999-09-28				
; PRIOR APPLICATION NUMBER: US 60/146,002				
; PRIOR FILING DATE: 1999-08-09				
; NUMBER OF SEQ ID NOS: 325720				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 97219				
; LENGTH: 1828				
; TYPE: DNA				
; ORGANISM: Human				
US-10-027-632-97219				
Query Match				
Best Local Similarity 67.7%; Score 17.6; DB 17; Length 1828;				
Pred. No. 1.9e+02;				
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				
Qy	1	TAAGGGTGTCCAGCAGCATCTGA	24	
Db	1439	TAAGGGTGTCCAGCAGCATCTGA	1416	
RESULT 10				
US-10-027-632-97220/c				
; Sequence 97220, Application US/10027632				
; Publication No. US20030204075A9				
; GENERAL INFORMATION:				
; APPLICANT: Wang, David G.				
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide				
; POLYMORPHISMS IN THE HUMAN GENOME				
; FILE REFERENCE: 10827.129				
; CURRENT APPLICATION NUMBER: US/10/027,632				
; CURRENT FILING DATE: 2002-04-30				
; PRIOR APPLICATION NUMBER: US 60/218,006				
; PRIOR FILING DATE: 2000-07-12				
; PRIOR APPLICATION NUMBER: US 60/198,676				
; PRIOR FILING DATE: 2000-04-20				
; PRIOR APPLICATION NUMBER: US 60/193,483				
; PRIOR FILING DATE: 2000-03-29				
; PRIOR APPLICATION NUMBER: US 60/185,218				
; PRIOR FILING DATE: 2000-02-24				
; PRIOR APPLICATION NUMBER: US 60/167,363				
; PRIOR FILING DATE: 1999-11-23				
; PRIOR APPLICATION NUMBER: US 60/156,358				
; PRIOR FILING DATE: 1999-09-28				
; PRIOR APPLICATION NUMBER: US 60/146,002				
; PRIOR FILING DATE: 1999-08-09				
; NUMBER OF SEQ ID NOS: 325720				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 97220				
; LENGTH: 1828				
; TYPE: DNA				
; ORGANISM: Oryza sativa				
FEATURE:				
OTHER INFORMATION: Clone ID: PAT_MRT4530_5227C.1				
US-10-027-632-97220				
Query Match				
Best Local Similarity 67.7%; Score 17.6; DB 18; Length 4354;				
Pred. No. 1.9e+02;				
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				
Qy	2	AAGGGGTTCGAGCAGCAACTGAC	25	
Db	1311	AAGGTCTTCGAGCAGCAACTGAC	1288	
RESULT 12				
US-09-864-761-29907/c				
; Sequence 29907, Application US/09864761				
; Patent No. US20020048763A1				
; GENERAL INFORMATION:				
; APPLICANT: Penn, Sharron G.				
; APPLICANT: Rank, David R.				
; APPLICANT: Hanzel, David K.				
; APPLICANT: Chen, Wensheng				
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR				
; GENE EXPRESSION ANALYSIS BY MICROARRAY				
; FILE REFERENCE: Aecmica-X-1				
; CURRENT APPLICATION NUMBER: US/09/864,761				
; CURRENT FILING DATE: 2001-05-23				
; PRIOR APPLICATION NUMBER: US 60/180,312				
; PRIOR FILING DATE: 2000-02-04				
; PRIOR APPLICATION NUMBER: US 60/207,456				
; PRIOR FILING DATE: 2000-05-26				
; PRIOR APPLICATION NUMBER: US 09/632,366				
; PRIOR FILING DATE: 2000-08-03				
; PRIOR APPLICATION NUMBER: GB 24263.6				
; PRIOR FILING DATE: 2000-10-04				
; PRIOR APPLICATION NUMBER: US 60/236,359				

```

, PRIOR APPLICATION NUMBER: US 60/236,359
, PRIOR FILING DATE: 2000-09-27
, PRIOR APPLICATION NUMBER: PCT/US01/00666
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00667
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00664
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00669
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00665
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00668
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00663
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00662
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00661
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00670
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/608,408
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/774,203
, PRIOR FILING DATE: 2001-01-29
, NUMBER OF SEQ ID NOS: 49117
, SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
, SEQ ID NO 13346
, LENGTH: 513
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
, OTHER INFORMATION: MAP TO AC010352.4
, OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.
, OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
US-09-864-761-13346

```

Query Match 66.2%; Score 17.2; DB 9; Length 513;
Best Local Similarity 86.4%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels

QY
4 GGGGTTCCAAGCCGCACTGAC 25

D_b
204 GGTTTCCGAGCGGCCACTGAC 183

```

RESULT 14
US-10-424-599-132352
; Sequence 132352, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132352
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(639)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:

```

```
; OTHER INFORMATION: Clone ID: PAT_WRT3847_90520C.1
US-10-424-599-132352

Query Match      66.2%; Score 17.2; DB 17; Length 639;
Best Local Similarity 86.4%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAAAGCGCGCAACTGAC 25
   |||||
Db 392 GGGGTTCCAAAGCGCGCGCAACTGAC 413

RESULT 15
US-09-770-445-47
; Sequence 47, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Huban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1191)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-445-47

Query Match      66.2%; Score 17.2; DB 9; Length 1191;
Best Local Similarity 86.4%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AGGGGTTCCAAAGCGCGCAACTGA 24
   |||||
Db 402 AGGGGTTCCAGGCTGCAACTCA 423

RESULT 16
US-10-723-860-6710/c
; Sequence 6710, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlornik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &

; OTHER INFORMATION: Clone ID: PAT_WRT3847_90520C.1
; FILE REFERENCE: 05882_0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6710
; LENGTH: 2339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-6710

Query Match      66.2%; Score 17.2; DB 18; Length 2339;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAAAGCGCGCAACTGAC 25
   |||||
Db 1440 GAGGTTCCAAAGTTGCAACTGAC 1419

RESULT 17
US-09-925-299-114
; Sequence 114, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 3731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3730)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-114

Query Match      66.2%; Score 17.2; DB 9; Length 3731;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGGTTCCAAAGCGCGCAACTGAC 25
   |||||
Db 396 GGTGTTCCGAGCGCGCAACTGAC 417

RESULT 18
US-09-925-299-114
; Sequence 114, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
```

```

Query Match      66.2%; Score 17.2; DB 17; Length 8906;
Best Local Similarity 86.4%; Pred. NO. 3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4   GGGGTTCCAAAGCGCAACTGAC 25
          |||||
DB     7128 GGTTTCGAGCGGCCACTGAC 7149

RESULT 21
US-10-182-936A-85
; Sequence 85, Application US/10182936A
; Publication NO. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-85

```

```

Query Match          66.2%; Score 17.2; DB 17; Length 8906;
Best Local Similarity 86.4%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 GGGGTTCCAGCGCCCAACTGAC 25
Db      7128 GGTGTTCCGAGCGCCCACTGAC 7149

RESULT 22
US-10-477-238A-664
; Sequence 664, Application US/10477238A
; Publication No. US20040221326A1
; GENERAL INFORMATION:
; APPLICANT: BabiJ, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 664
; LENGTH: 8906
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-477-238A-664

Query Match          66.2%; Score 17.2; DB 18; Length 8906;
Best Local Similarity 86.4%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 GGGGTTCCAGCGCCCAACTGAC 25
Db      7128 GGTGTTCCGAGCGCCCACTGAC 7149

RESULT 23
US-10-680-287A-664
; Sequence 664, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: BabiJ, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; CURRENT FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 664

```



```
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13237C.1
US-10-437-963-6558

Query Match      65.4%; Score 17; DB 18; Length 280;
Best Local Similarity 80.0%; Pred.No. 3.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      2 AAGGGGTTCCCAAGCGCGCAACTGACG 26
      ||||| ||||| ||||| ||||| |||||
Db      165 AAGGAGCTCCCAAGTCCCATCTGACG 189
```

```
Search completed: March 12, 2005, 00:25:15
Job time : 130.89 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 344.049 Seconds
(without alignments)
2876.537 Million cell updates/sec

Title: US-09-674-277-11
Perfect score: 26
Sequence: 1 taaggggttcacagccgaactgacg 26

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution...

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.4	74.6	748	7	CV064706
2	19.2	73.8	955	7	W29474 mb99f10.r1
3	18.8	72.3	337	5	B0216934 603758737
4	18.8	72.3	821	4	B1872872 603397485
5	18.8	72.3	824	9	CC905643 t024i14ba
6	18.6	71.5	174	1	AA190947 zp83f03.r
7	18.6	71.5	350	6	CB781359 AMGNNUC:C
8	18.6	71.5	505	5	BQ825787 1030130A0
9	18.6	71.5	502	5	BQ825786 1030130A0
10	18.6	71.5	573	4	BI951276 HVSME1002
11	18.6	71.5	615	4	BG393915 602416790
12	18.6	71.5	643	2	BE547564 601075308
13	18.6	71.5	645	8	B2278356 CH230-392
14	18.6	71.5	721	2	BF529080 602040090
15	18.6	71.5	742	6	CD778991 EST650352
16	18.6	71.5	765	9	CC924307 t074a06ba
17	18.6	71.5	766	2	BE563890 601348101
18	18.6	71.5	870	9	CR072040 Forward s
19	18.6	71.5	905	6	CD778990 EST650351
20	18.6	71.5	1447	2	BE962423 601655714
21	18.4	70.8	903	4	BI093748 602860456
22	18.2	70.0	378	9	CR123873 Reverse s
23	18.2	70.0	791	9	CL936731 OA_Aba004
24	18.2	70.0	965	5	BQ925391 AGENC007

C 25	18.2	70.0	1038	9	CL035478	CL035478	CH216-39N
C 26	18	69.2	412	7	CK468268	CK468268	939646 MA
C 27	18	69.2	545	7	CN989004	CN989004	64963 125
C 28	18	69.2	547	7	CN986469	CN986469	62086 125
C 29	18	69.2	548	8	B2899328	B2899328	CH240-15L
C 30	18	69.2	565	2	BE368858	BE368858	601221488
C 31	18	69.2	568	7	CN978789	CN978789	29933-125
C 32	18	69.2	589	7	CN990905	CN990905	67176-125
C 33	18	69.2	630	7	CN979150	CN979150	30337-125
C 34	18	69.2	653	4	BJ232437	BJ232437	BJ232437
C 35	18	69.2	710	7	CN976117	CN976117	26041-125
C 36	18	69.2	716	7	CN974680	CN974680	23462-125
C 37	18	69.2	716	7	CN978621	CN978621	30888-125
C 38	18	69.2	717	7	CN989184	CN989184	65178-125
C 39	18	69.2	724	7	CN987591	CN987591	63311-125
C 40	18	69.2	747	7	CN987312	CN987312	63005-125
C 41	18	69.2	751	9	CR796337	CR796337	GROAA14B
C 42	18	69.2	757	7	CN976325	CN976325	26374-125
C 43	18	69.2	781	5	BP141859	BP141859	BP141859
C 44	18	69.2	800	7	CK466163	CK466163	937342 MA
C 45	18	69.2	811	7	CN992486	CN992486	70429-125

ALIGNMENTS

RESULT 1
LOCUS CV064706
DEFINITION WNEL14b3 Wheat EST endosperm library Triticum aestivum cDNA clone
WNEL14b3 5' similar to Oryza sativa, mRNA sequence.
ACCESSION CV064706
VERSION
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 748)
AUTHORS Ali, S., Holloway, B. and Taylor, W.C.
TITLE Normalisation of cereal endosperm EST libraries for structural and functional genomic analysis
JOURNAL Plant Mol. Biol. Rep. 18, 123-132 (2000)
COMMENT Contact: Bill Taylor
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industry,
CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia
Tel: 61 2 6246 5223
Fax: 61 2 6246 5000
Email: Bill.Taylor@csiro.au
Seg primer: M13 reverse primer
High quality sequence stop: 748.
Location/Qualifiers
1. .748
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Hartog"
/db_xref="taxon:4565"
/clone="WNEL14b3"
/tissue_type="endosperm"
/dev_stage="developing endosperm tissue 6, 8, 10 dpa (days post anthesis)"
/lab_host="DH10B (Life Technology)"
/note="Vector: _ziplox; Site: Sal I; Site 2: Not I; mRNA was prepared from endosperm tissues of the wheat cultivar Hartog. cDNA was synthesised from pooled 6, 8, and 10 dpa endosperm using Not I-oligo(dT)18 primer/adaptor (Pharmacia Biotech), and then ligated to the Sal I-Not I site of _ziplox vector (Life Technology) after adding a Sal I-Xho I adaptor (Stratagene). Constructed by Shahjahan Ali and Bill Taylor."

ORIGIN

Query Match 74.6%; Score 19.4; DB 7; Length 748;
 Best Local Similarity 95.2%; Pred. No. 1.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TAAGGGGTTCACAGCGCAAC 21
 |||||
 Db 622 TAAGGGGTTCACAGCGCAATC 642

RESULT 2

W29474/c
 LOCUS W29474 955 bp mRNA linear EST 11-SEP-1996
 DEFINITION m599f10.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone for
 IMAGE:337579 5' similar to gp:X74856 M.musculus L28 mRNA for
 ribosomal protein L28 (MOUSE);, mRNA sequence.

ACCESSION W29474.1 GI:1309619

VERSION EST.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 955)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:218979

Seq primer: ETPPrimer

High quality sequence stop: 427.

Location/Qualifiers

FEATURES

source

1..955
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:337579"
 /dev_stage="19.5 dpc total fetus"
 /lab_hosts="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NMF19.5"
 /note="Vector: pT7T3D (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."

ORIGIN

Query Match 73.8%; Score 19.2; DB 7; Length 955;
 Best Local Similarity 87.5%; Pred. No. 2.5e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAGGGGTTCACAGCGCAACTGAC 25
 |||||
 Db 627 AAGGGGTTCACAGCAAAATTGAC 604

RESULT 3

BU216934/c

LOCUS BU216934 337 bp mRNA linear EST 25-NOV-2002

DEFINITION 603758737F1 CSBQCHN04 Gallus gallus cDNA clone CHES1673j9 5', mRNA

sequence.

ACCESSION BU216934

VERSION BU216934.1 GI:25397621

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 337)

AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.O.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE JOURNAL

MEDLINE 22335534

COMMENT 12443392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..337

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHES1673j9"

/tissue_type="whole embryo"

/dev_stage="20-21"

/lab_hosts="DH10B"

/clone_lib="CSBQCHN04"

/note="Organ: whole embryo; Vector: pBluescript II KS(+);

Site 1: EcoRI; Site 2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunted, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Query Match 72.3%; Score 18.8; DB 5; Length 337;
 Best Local Similarity 90.9%; Pred. No. 3.5e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AGGGGTTCACAGCGCAACTGA 24
 |||||

Db 25 AGGTGTTCCACAGCAACTGA 4

RESULT 4

BI872872

LOCUS BI872872 821 bp mRNA linear EST 11-OCT-2001

DEFINITION 603397485F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5401076 5',

mRNA sequence.

ACCESSION BI872872

VERSION BI872872.1 GI:16046547

```

KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM

REFERENCE
1 (bases 1 to 821)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12022 row: j column: 21
            High quality sequence stop: 690.

FEATURES
source
1..821
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5401076"
/lab_host="DH10B (phage-resistant)"
/tissue_type="retina"
/clone_lib="NIH MGC 94"
note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match      72.3%; Score 18.8; DB 4; Length 821;
Best Local Similarity 90.9%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAGGGTTCCAAAGCGCAACTG 23
|||||
Db 735 AAGGGTTCCAAAGCGCAACTG 756

RESULT 5
CC905643/c
LOCUS      824 bp DNA linear GSS 08-AUG-2003
DEFINITION t024j14ba.r1 TAMBT Bos taurus genomic clone t024j14ba, genomic
survey sequence.
ACCESSION  CC905643
VERSION     CC905643.1 GI:33524576
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 824)
Lin, S., Najjar, F. Z., Adelson, D., Gill, C. A. and Roe, B. A.
Bovine BAC End Sequences from Library TAMBT
Unpublished (2003)
Contact: Bruce A. Roe
Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Class: BAC ends
High quality sequence start: 43
High quality sequence stop: 355.

FEATURES
Location/Qualifiers

```

```

source
1..824
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Angus bull T A M U Shoshone Y6 11519666"
/db_xref="taxon:9913"
/clone="t024j14ba"
/sex="Male"
/cell_type="Blood"
/clone_lib="TAMBT"
/note="Vector: pBeloBAC11; Site 1: HindIII; Site 2:
HindIII; TAMBT Bovine BAC library (Male) produced by Texas
A&M University, Department of Animal Science."

ORIGIN
Query Match      72.3%; Score 18.8; DB 9; Length 824;
Best Local Similarity 90.9%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAGGGTTCCAAAGCGCAACTG 23
|||||
Db 577 AAGGGTTCCCTCGCGCAACTG 556

RESULT 6
AA190947/c
LOCUS      174 bp mRNA linear EST 10-MAR-1998
DEFINITION zp83f03.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
IMAGE:626813 5', mRNA sequence.
ACCESSION  AA190947
VERSION     AA190947.1 GI:1779469
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 174)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1860 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 116.

FEATURES
source
1..174
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5045511"
/db_xref="taxon:9606"
/clone="IMAGE:626813"
/sex="female"
/dev_stage="HeLa S3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene HeLa cell s3 937216"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. HeLa S3
epitheloid carcinoma cells grown to semi-confluency
without induction. Average insert size: 1.5 kb; Uni-ZAP XR
Vector. ~5' adaptor sequence: 5' GAATTCGCGCAG 3' ~3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
Query Match      71.5%; Score 18.6; DB 1; Length 174;

```

Best Local Similarity 84.0%; Pred. No. 4.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAAGGGGTTCCAGCGCAACTGAC 25
||||| ||||| ||||| ||||| |||||
Db 73 TAAGGGGTTTCAGCGCGCATCTGAC 49

RESULT 7
CB781359/c
LOCUS
DEFINITION AMGNNUC:CR1-00004-E3-A Colon Rat 1 (10390) Rattus norvegicus CDNA
Clone cr1-00004-e3 5', mRNA sequence.

ACCESSION CB781359
VERSION CB781359.1 GI:29869750
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 350)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00004 row: e column: 3.

FEATURES
source
1..350
Location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="cr1-00004-e3"
/tissue_type="intestine, fetal"
/clone_lib="Colon Rat 1 (10390)"
/note=vector: pMOB; Site_1: SalI; Site_2: NotI; Colon Rat
1 Fetal rat intestinal cDNA library."

ORIGIN

Query Match 71.5%; Score 18.6; DB 6; Length 350;
Best Local Similarity 84.0%; Pred. No. 4.4e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AAGGGGTTCCAGCGCAACTGACG 26
||||| ||||| ||||| ||||| |||||
Db 255 AAGAGGTTCCATCAGCGCAACTGACG 231

RESULT 8
BQ825787/c
LOCUS
DEFINITION BQ825787 502 bp mRNA linear EST 01-AUG-2002
1030130A06.y1 C. reinhardtii CC-1690, Deflagellation (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BQ825787
VERSION BQ825787.1 GI:22077011
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii

ORGANISM Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaeae; Chlamydomonas.

REFERENCE 1 (bases 1 to 502)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
TITLE Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1030

JOURNAL Unpublished (2002)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University

Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

FEATURES

source

1..502
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Deflagellation
(normalized), Lambda Zap II"
/note=vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Deflagellation library, constructed by John Davies
and Jeffrey McDermott, combines cDNAs from CC-1690 cells
which had been re-synthesizing flagella for 15, 30 and 60
min after being deflagellated by pH shock. PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with Exassist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al., (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 71.5%; Score 18.6; DB 5; Length 502;
Best Local Similarity 84.0%; Pred. No. 4.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAAGGGGTTCCAGCGCAACTGAC 25
||||| ||||| ||||| ||||| |||||
Db 461 TAAGGGGTTTCAGCGCGCATCTGAC 437

RESULT 9

BQ825786

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..505

Location/Qualifiers

/organism="Chlamydomonas reinhardtii"

/mol_type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_lib="C. reinhardtii CC-1690, Deflagellation
(normalized), Lambda Zap II"/note=vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Deflagellation library, constructed by John Davies
and Jeffrey McDermott, combines cDNAs from CC-1690 cells
which had been re-synthesizing flagella for 15, 30 and 60


```

Qy 1 TAAGGGTTCCAGCCGCACTGAC 25
    ||||| ||||| ||||| ||||| |||||
Db 124 TAAGGGTTCCAGCCGCCCCGAC 100

RESULT 12
BE547564/c
LOCUS
DEFINITION 601075308F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3461592 5',
            mRNA sequence.
ACCESSION BE547564
VERSION BE547564.1 GI:9776209
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 643)
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: Incyte Genomics, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LHAM8457 row: o column: 01
          High quality sequence stop: 624.
FEATURES
          source
            1..643
              Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:3461592"
                /tissue_types="cervical carcinoma cell line"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_12"
                /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: Not I;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.4 Kb. Library prepared by Life
                Technologies."
ORIGIN
          Query Match 71.5%; Score 18.6; DB 2; Length 643;
          Best Local Similarity 84.0%; Pred. No. 4.7e+02;
          Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAAGGGTTCCAGCCGCACTGAC 25
    ||||| ||||| ||||| ||||| |||||
Db 109 TAAGGGTTCCAGCCGCCCCGAC 85

RESULT 13
BZ278356
LOCUS
DEFINITION CH230-392P5.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
            CH230-392P5, genomic survey sequence.
ACCESSION BZ278356
VERSION BZ278356.1 GI:24003055
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 645)
          Zhao, S., Shetty, J., Shatsman, S., Tsengay, G., Geer, K.,
          Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,

```

```

Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other_GSSs: CH230-392P5.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 392 row: P column: 5
Seq primer: T7
Class: BAC ends.
FEATURES
          source
            1..645
              Location/Qualifiers
                /organism="Rattus norvegicus"
                /mol_type="genomic DNA"
                /strain="BN/SsNHsd/MCW"
                /db_xref="taxon:10116"
                /clone="CH230-392P5"
                /sex="Female"
                /cell_type="Brain"
                /clone_lib="CHORI-230 Segment 2"
                /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
                CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                Pieter de Jong"
ORIGIN
          Query Match 71.5%; Score 18.6; DB 8; Length 645;
          Best Local Similarity 84.0%; Pred. No. 4.7e+02;
          Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAAGGGTTCCAGCCGCACTGAC 25
    ||||| ||||| ||||| ||||| |||||
Db 530 TAAGAGGTTCCACCACTGAC 554

RESULT 14
BF529080/c
LOCUS
DEFINITION 602040090F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4177870
            5', mRNA sequence.
ACCESSION BF529080
VERSION BF529080.1 GI:11616537
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 721)
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: David N. Louis, M.D.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LHAM9485 row: c column: 23
          High quality sequence stop: 609.
FEATURES
          source
            1..721

```



```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4177870"
/tissue_type="anaplastic oligodendroglioma with lp/19q loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 71.5%; Score 18.6; DB 2; Length 721;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TAAGGGTTCCAAAGCGCAACTGAC 25
Db 152 TAAGGGTTCCAAAGCGCGCCCGAC 128

RESULT 15
LOCUS CD778991 742 bp mRNA linear EST 01-JUL-2004
DEFINITION EST650352 RAA Rhipicephalus appendiculatus cDNA clone RAA562 5' end, mRNA sequence.
ACCESSION CD778991
VERSION CD778991.1 GI:49534664
KEYWORDS EST.
SOURCE Rhipicephalus appendiculatus
ORGANISM Rhipicephalus appendiculatus
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari; Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Rhipicephalus.
REFERENCE 1 (bases 1 to 742)
AUTHORS Nene,V., Lee,Y., Skilton,R., Mwaura,S., Quackenbush,J., Gardner,M. and Bishop,R.
TITLE An index of genes transcribed in the salivary glands of Rhipicephalus appendiculatus
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST650351
Contact: Vishvanath Nene
Parasite Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-610-5968
Fax: 301-838-0208
Email: nene@igr.org

Seq primer: M13 reverse.
Location/Qualifiers
1..742
/organism="Rhipicephalus appendiculatus"
/mol_type="mRNA"
/db_xref="taxon:34631"
/clone="RAA562"
/dev_stage="Adult"
/lab_host="E. coli strain DH10B-Tona"
/clone_lib="RAA"
/note="Organ: Salivary glands; Vector: pCMVSPORT6.0.ccdB; Salivary Glands were dissected on day four after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli. Library RAA was made from uninfected ticks."

ORIGIN
Query Match 71.5%; Score 18.6; DB 6; Length 742;

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4177870"
/tissue_type="anaplastic oligodendroglioma with lp/19q loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 71.5%; Score 18.6; DB 2; Length 721;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TAAGGGTTCCAAAGCGCAACTGAC 25
Db 152 TAAGGGTTCCAAAGCGCGCCCGAC 128

RESULT 15
LOCUS CD778991 742 bp mRNA linear EST 01-JUL-2004
DEFINITION EST650352 RAA Rhipicephalus appendiculatus cDNA clone RAA562 5' end, mRNA sequence.
ACCESSION CD778991
VERSION CD778991.1 GI:49534664
KEYWORDS EST.
SOURCE Rhipicephalus appendiculatus
ORGANISM Rhipicephalus appendiculatus
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari; Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Rhipicephalus.
REFERENCE 1 (bases 1 to 742)
AUTHORS Nene,V., Lee,Y., Skilton,R., Mwaura,S., Quackenbush,J., Gardner,M. and Bishop,R.
TITLE An index of genes transcribed in the salivary glands of Rhipicephalus appendiculatus
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST650351
Contact: Vishvanath Nene
Parasite Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-610-5968
Fax: 301-838-0208
Email: nene@igr.org

Seq primer: M13 reverse.
Location/Qualifiers
1..742
/organism="Rhipicephalus appendiculatus"
/mol_type="mRNA"
/db_xref="taxon:34631"
/clone="RAA562"
/dev_stage="Adult"
/lab_host="E. coli strain DH10B-Tona"
/clone_lib="RAA"
/note="Organ: Salivary glands; Vector: pCMVSPORT6.0.ccdB; Salivary Glands were dissected on day four after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli. Library RAA was made from uninfected ticks."

ORIGIN
Query Match 71.5%; Score 18.6; DB 6; Length 742;

Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAGGGTTCCAAAGCGCAACTGAC 26
Db 576 AAGGGTTCCAAAGCGCGCCCGAC 600

RESULT 16
LOCUS CC924307 765 bp DNA linear GSS 08-AUG-2003
DEFINITION CC924307.1 GI:33559646
ACCESSION CC924307
VERSION CC924307.1 GI:33559646
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 765)
AUTHORS Lin,S., Najjar,F.Z., Adelson,D., Gill,C.A. and Roe,B.A.
TITLE Bovine BAC End Sequences from Library TAMBT
JOURNAL Unpublished (2003)
COMMENT Contact: Bruce A. Roe
Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Class: BAC ends
High quality sequence start: 33
High quality sequence stop: 687.

FEATURES
Location/Qualifiers
1..765
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="t074a06ba"
/sex="Male"
/cell_type="Blood"
/clone_lib="TAMBT"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2: HindIII; TAMBT Bovine BAC library (Male) produced by Texas A&M University, Department of Animal Science."

ORIGIN
Query Match 71.5%; Score 18.6; DB 9; Length 765;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TAAGGGTTCCAAAGCGCAACTGAC 25
Db 669 TAAGGGTTTCAGAGTCGCACTAAC 693

RESULT 17
LOCUS BE563890 766 bp mRNA linear EST 15-AUG-2000
DEFINITION BE563890.1 GI:9807610
ACCESSION BE563890
VERSION BE563890.1 GI:9807610
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 766)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
```

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM376 row: b column: 01 High quality sequence stop: 554.

ORIGIN

Query Match	71.5%	Score 18.6;	DB 2;	Length 766;
Best Local Similarity	84.0%;	Pred. No. 4.8e+02;		
Matches 21: Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY 1 TAAGGGTTCCAAGCCGCACTGAC 25
|||
Db 128 TAAGGGTTTCCAAGCGCCCCGAC 104
|||

RESULT 18	CR072040/c	CR072040	870 bp	DNA	linear	GSS 05-JUL-2004
LOCUS						
DEFINITION		Forward strand read from insert in 3'HPT insertion targeting and chromosome engineering clone MHP168i21, genomic survey sequence.				
ACCESSION		CR072040				
VERSION		CR072040.1	GI:49805630			
KEYWORDS		GSS: genome survey sequence; MICR.				

REFERENCE
AUTHORS
1 (Bases 1 to 870)
Adams, D. J., Biggs, P. J., Cox, A. V., Davies, R. M., van der Weyden, L.,
Jongers, J., Smith, J., Plumb, R. W., Taylor, R. G., Nishijima, I., Yu, Y.,
Rogers, J. and Bradley, A.
TITLE
Direct Submission

FEATURES
source

RESULT 20	BE962423	1447 bp	mRNA	linear	EST 14-DEC-2000
LOCUS	601655714R1	NIH_MGC_65	Homo sapiens	cDNA clone	IMAGE:3846290 3',
DEFINITION	mRNA sequence.				
ACCESSION	BE962423				

```

VERSION      BE962423.2  GI:11765310
KEYWORDS
SOURCE       EST.
ORGANISM     Homo sapiens (human)
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        1 (bases 1 to 1447)
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              On Oct 3, 2000 this sequence version replaced gi:10573128.
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L10573128
              High quality sequence stop: 279.

FEATURES     Location/Qualifiers
             1..1447
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:3846290"
             /tissue_type="adenocarcinoma"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_65"
             /note="Organ: colon; Vector: pCMV-Sport6; Site: 1: NotI;
             Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
             Average insert size 1.8 kb. Library constructed by Life
             Technologies."
ORIGIN
Query Match      71.5%; Score 18.6; DB 2; Length 1447;
Best Local Similarity 84.0%; Pred. No. 5.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      2  AAGGGTTCCTCAAGCGCAACTGAC 26
        |||||
        952 AAGGGTTCCTCAAGCGCAACTGAC 976

Db
RESULT 21
BI093748
LOCUS
DEFINITION      903 bp mRNA linear EST 20-JUN-2001
VERSION          BI093748.1  GI:14512078
KEYWORDS
SOURCE           Homo sapiens (human)
ORGANISM         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1 (bases 1 to 903)
AUTHORS          NIH-MGC http://mgi.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL          Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  cDNA Library Preparation: Life Technologies, Inc.
                  cDNA Library Arrayed by: Incyte Genomics, Inc.
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L10573128
                  High quality sequence stop: 661.

FEATURES     Location/Qualifiers
             1..903
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5001829"
             /cell_line="MGC36"
             /lab_host="DH10B"
             /clone_lib="NIH_MGC_10"
             /note="Organ: cervix; Vector: pCMV-Sport6; Site 1: NotI;
             Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
             Average insert size 1.5 kb. Library prepared by Life
             Technologies."
ORIGIN
Query Match      70.0%; Score 18.4; DB 4; Length 903;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      6  GGTTCCTCAAGCGCAACTGAC 25
        |||||
        822 GGTTCCTCAAGCGCAACTGAC 841

Db
RESULT 22
CR123873/c
LOCUS
DEFINITION      378 bp DNA linear GSS 06-JUL-2004
VERSION          CR123873.1  GI:49871325
KEYWORDS         GSS; genome survey sequence; MICER.
SOURCE           Mus musculus
ORGANISM         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE        1 (bases 1 to 378)
AUTHORS          Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
                  Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
                  Rogers,J. and Bradley,A.
TITLE            Direct Submission
JOURNAL          Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
                  CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES         Location/Qualifiers
             1..378
             /organism="Mus musculus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10090"
             /clone="MHP92p17"
             /clone_lib="MHPp"
ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 378;
Best Local Similarity 87.0%; Pred. No. 6.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      4  GGGTTCCTCAAGCGCAACTGAC 26
        |||||
        36 GGGTTCCTCAAGCGCAACTGAC 14

Db
RESULT 23
CL936731
LOCUS
DEFINITION      791 bp DNA linear GSS 14-SEP-2004
VERSION          OA_ABA0047118.5  genomic survey sequence.
KEYWORDS         OA_ABA0047118 5', genomic survey sequence.
SOURCE           CL936731.1  GI:52067562
ORGANISM         Oryza australiensis
                  Oryza australiensis
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

```

REFERENCE
AUTHORS      Ehrhartoideae; Oryzae; Oryza.
              1 (bases 1 to 791)
JOURNAL      Kim, H., Yu, Y., Stum, D., Yost, D., Rao, K., Luo, M., Jetty, R.,
COMMENT      Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.
              OMAP Project
              Unpublished (2004)
              Contact: Rod A. Wing
              Arizona Genomics Institute
              University of Arizona
              Forbes Building Room 303, Tucson, AZ 85721-0036, USA
              Tel: 520 626 9595
              Fax: 520 621 1259
              Email: http://genome.arizona.edu
              PCR Primers
              FORWARD: TAA TAC GAC TCA CTA TAG GG
              BACKWARD: CAC TCA TTA GGC ACC CCA
              Plate: 0047 row: I column: 18
              Seq primer: TAA TAC GAC TCA CTA TAG GG
              Class: BAC ends.

FEATURES
source
    location/Qualifiers
    1..791
    /organism="Oryza australiensis"
    /mol_type="genomic DNA"
    /db_xref="taxon:4532"
    /clone="OA_ABA0047118"
    /tissue_type="young leaves"
    /lab_host="DH10B T1 phage resistant"
    /clone_lib="OA ABA"
    /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 791;
Best Local Similarity 87.0%; Pred. No. 7.6e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      4 GGGGTTCCAGCCGCAACTGAC 26
      |||||||
Db      370 GGGGTTCCAGCCGCAACTCCG 392

RESULT 24
LOCUS      BQ925391
DEFINITION AGENCOURT_8763855 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6327841
            5', mRNA sequence.
ACCESSION  BQ925391
VERSION    BQ925391.1 GI:22340422
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 965)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
            cDNA Library Preparation: ResGen, Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM13778 row: b column: 02
            High quality sequence start: 11
            High quality sequence stop: 592.
            Location/Qualifiers
            1..965
            /organism="Mus musculus"
            /mol_type="mRNA"

REFERENCE
AUTHORS      BQ925391
JOURNAL      BQ925391
COMMENT      BQ925391.1 GI:22340422

FEATURES
source
    location/Qualifiers
    1..1038
    /organism="Xenopus tropicalis"
    /mol_type="genomic DNA"
    /strain="Nigerian frog"
    /db_xref="taxon:8364"
    /clone="CH216-39N15"
    /sex="male"
    /cell_line="Stock 248 F7A2, inbred N7"
    /clone_lib="CH216"
    /note="Vector: pPARBAC2.1; CHORI-216 Xenopus tropicalis
    BAC library"

ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 1038;
Best Local Similarity 87.0%; Pred. No. 7.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 AAGGGTTCCAGCCGCAACTGA 24
      |||||||
Db      979 AAGGGTTCCAGCCGCAATTGA 957

Search completed: March 11, 2005, 13:01:07
Job time : 346.049 secs

```

```

/db_xref="taxon:10090"
/clone="IMAGE:6327841"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_130"
/note="Organ: otcysts; Vector: pCMV-SPORT6.1; Site_1:
EcorV; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.95 kb. Constructed by
ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      70.0%; Score 18.2; DB 5; Length 965;
Best Local Similarity 87.0%; Pred. No. 7.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 AAGGGTTCCAGCCGCAACTGA 24
      |||||||
Db      856 AAGGGTTCCAGCCGCAACTGA 878

```

```

RESULT 25
LOCUS      CL035478
DEFINITION CH216-39N15_RML1.1 CH216 Xenopus tropicalis genomic clone
            CH216-39N15, genomic survey sequence.
ACCESSION  CL035478
VERSION    CL035478.1 GI:40488526
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
            Xenopodinae; Xenopus; Silurana.
            1 (bases 1 to 1039)
            Krematzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@wustl.edu
            Insert Length: 175000 Std Error: 0.00
            Seq primer: RML1 TAGCACTACTATAGGAGA
            Class: BAC ends
            High quality sequence start: 53
            High quality sequence stop: 844.
            Location/Qualifiers
            1..1038
            /organism="Xenopus tropicalis"
            /mol_type="genomic DNA"
            /strain="Nigerian frog"
            /db_xref="taxon:8364"
            /clone="CH216-39N15"
            /sex="male"
            /cell_line="Stock 248 F7A2, inbred N7"
            /clone_lib="CH216"
            /note="Vector: pPARBAC2.1; CHORI-216 Xenopus tropicalis
            BAC library"

ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 1038;
Best Local Similarity 87.0%; Pred. No. 7.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 AAGGGTTCCAGCCGCAACTGA 24
      |||||||
Db      979 AAGGGTTCCAGCCGCAATTGA 957

```

```

RESULT 25
LOCUS      CL035478/c
DEFINITION CH216-39N15_RML1.1 CH216 Xenopus tropicalis genomic clone
            CH216-39N15, genomic survey sequence.
ACCESSION  CL035478
VERSION    CL035478.1 GI:40488526
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
            Xenopodinae; Xenopus; Silurana.
            1 (bases 1 to 1039)
            Krematzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@wustl.edu
            Insert Length: 175000 Std Error: 0.00
            Seq primer: RML1 TAGCACTACTATAGGAGA
            Class: BAC ends
            High quality sequence start: 53
            High quality sequence stop: 844.
            Location/Qualifiers
            1..1038
            /organism="Xenopus tropicalis"
            /mol_type="genomic DNA"
            /strain="Nigerian frog"
            /db_xref="taxon:8364"
            /clone="CH216-39N15"
            /sex="male"
            /cell_line="Stock 248 F7A2, inbred N7"
            /clone_lib="CH216"
            /note="Vector: pPARBAC2.1; CHORI-216 Xenopus tropicalis
            BAC library"

ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 1038;
Best Local Similarity 87.0%; Pred. No. 7.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 AAGGGTTCCAGCCGCAACTGA 24
      |||||||
Db      979 AAGGGTTCCAGCCGCAATTGA 957

```

```

Query Match      70.0%; Score 18.2; DB 9; Length 791;
Best Local Similarity 87.0%; Pred. No. 7.6e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      4 GGGGTTCCAGCCGCAACTGAC 26
      |||||||
Db      370 GGGGTTCCAGCCGCAACTCCG 392

```

```

RESULT 24
LOCUS      BQ925391
DEFINITION AGENCOURT_8763855 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6327841
            5', mRNA sequence.
ACCESSION  BQ925391
VERSION    BQ925391.1 GI:22340422
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 965)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
            cDNA Library Preparation: ResGen, Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM13778 row: b column: 02
            High quality sequence start: 11
            High quality sequence stop: 592.
            Location/Qualifiers
            1..965
            /organism="Mus musculus"
            /mol_type="mRNA"

REFERENCE
AUTHORS      BQ925391
JOURNAL      BQ925391
COMMENT      BQ925391.1 GI:22340422

FEATURES
source
    location/Qualifiers
    1..1038
    /organism="Xenopus tropicalis"
    /mol_type="genomic DNA"
    /strain="Nigerian frog"
    /db_xref="taxon:8364"
    /clone="CH216-39N15"
    /sex="male"
    /cell_line="Stock 248 F7A2, inbred N7"
    /clone_lib="CH216"
    /note="Vector: pPARBAC2.1; CHORI-216 Xenopus tropicalis
    BAC library"

ORIGIN
Query Match      70.0%; Score 18.2; DB 9; Length 1038;
Best Local Similarity 87.0%; Pred. No. 7.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 AAGGGTTCCAGCCGCAACTGA 24
      |||||||
Db      979 AAGGGTTCCAGCCGCAATTGA 957

```

```

Search completed: March 11, 2005, 13:01:07
Job time : 346.049 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 45.3508 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-12
Perfect score: 31
Sequence: 1 ctcaaggcagtcgagtcggttggaac 31

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	31	3 AAZ36112	Aaz36112 Primer de
2	31	100.0	1489	3 AAZ36101	Aaz36101 Nucleic a
3	27	87.1	31	3 AAZ36113	Aaz36113 Primer de
C 4	21	67.7	26	3 AAZ36110	Aaz36110 Primer de
C 5	20.2	65.2	664	10 ADB68745	Adb68745 Pseudomon
C 6	20	64.5	26	3 AAZ36111	Aaz36111 Primer de
C 7	20	64.5	636	6 AAL49962	Aal49962 Human B11
8	20	64.5	824	6 AAL49970	Aal49970 Human B11
9	20	64.5	1018	6 AAV77583	Avv77583 Glycerald
10	20	64.5	1143	9 ADB80402	Adb80402 Human MDD
C 11	20	64.5	1770	12 ADL03792	Adl03792 DNA encod
12	20	64.5	1858	10 ADA53102	Ada53102 Human cod
13	20	64.5	3596	6 AAL49947	Aal49947 Human B11
14	20	64.5	3614	12 ADQ64982	Adq64982 Novel hum
C 15	20	64.5	23210	4 AAF28530	Aaf28530 Genomic F
16	19.8	63.9	3237	13 ADS63488	Ads63488 Bacterial
17	19.8	63.9	3240	13 ADS64237	Ads64237 Bacterial
18	19.8	63.9	3240	13 ADS63858	Ads63858 Bacterial
19	19.4	62.6	242	12 ACH83759	Ach83759 Human gen
20	19.4	62.6	509	12 ACH70059	Ach70059 Human gen

C 21	19.4	62.6	1391	3 AAC34842	Aac34842 Arabidops
22	19.4	62.6	2700	4 AAD18767	Aad18767 Human ANI
23	19.4	62.6	3017	13 ADR58978	Adr58978 Human Elk
24	19.4	62.6	6862	13 ADR58980	Adr58980 Human Elk
25	19.4	62.6	6908	13 ADR59004	Adr59004 Human Elk
26	19.4	62.6	41626	13 ADS36500	Ads36500 Human aut
C 27	19.2	61.9	1156	8 ACA48626	Aca48626 Prokaryot
C 28	19.2	61.9	1278	12 ADO03059	Ado03059 Corn orth
C 29	19.2	61.9	1278	13 ADI42860	Adi42860 Plant tra
C 30	19.2	61.9	1362	13 ADT746111	Adt746111 Bacterial
C 31	19.2	61.9	1662	8 ACA50970	Aca50970 Prokaryot
C 32	19.2	61.9	1686	8 ACA51630	Aca51630 Prokaryot
C 33	19	61.3	588	7 ADR41303	Adr41303 Human CD-
C 34	19	61.3	724	4 AAS26853	Aas26853 Human CDN
C 35	19	61.3	1584	8 ABZ76273	Abz76273 Human sol
C 36	19	61.3	1632	13 ADS60246	Ads60246 Bacterial
C 37	19	61.3	2052	8 AAD52530	Aad52530 Human not
C 38	19	61.3	2053	8 ABX75303	Abx75303 Human not
C 39	19	61.3	2053	8 ABZ81800	Abz81800 Human not
C 40	19	61.3	2055	2 AAV68524	Aav68524 Nucleotid
C 41	19	61.3	2055	3 AAZ60905	Aaz60905 DNA encod
C 42	19	61.3	2055	10 AAD65064	Aad65064 Human del
C 43	19	61.3	2058	3 AAZ91643	Aaz91643 Human not
C 44	19	61.3	2058	13 ADR88264	Adr88264 Human Del
C 45	19	61.3	2058	13 ADS14396	Ads14396 Human Del

ALIGNMENTS

RESULT 1

AAZ36112
ID AAZ36112 standard; DNA; 31 BP.

XX AAZ36112;

DT 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

PN WO9955908-A2.

XX WO9955908-A2.

PD 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

PT Escherichia coli, particularly serotype O157:H7, used for detecting these

PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

CC AAZ36103-27 represent fragments derived from nucleic acid sequences specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are derived from two sequences. The first (AAZ36101) is 99.9% homologous to the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).

CC The second sequence (AAZ36102) is associated with the presence of

```

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
XX
SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTCGGCTTGAAC 31
DB 1 CTCACGGCATCGTCAGTTCGGCTTGAAC 31

RESULT 2
AAZ36101
ID AAZ36101 standard; DNA; 1489 BP.
XX
AC AAZ36101;
XX
DT 11-FEB-2000 (first entry)
XX
DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; ds.
XX
OS Escherichia coli.
XX
PN WO9555908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
PA (SNFI ) PASTEUR SANOPI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
WPI; 2000-013443/01.
XX
New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 1; Fig 1; 48pp; French.
XX
The present sequence is specific to enterohemorrhagic Escherichia coli
CC (EHEC). The sequence is 99.9% homologous to the katP gene of E. coli
CC O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
CC homologous with IS91 of E. coli (nucleotides 1-406 of the present
CC sequence). The present sequence is of plasmid origin. Fragments of the
CC present sequence are used as probes and primers for detection of E.
CC coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
CC animal samples, foods or the environment. The fragments are also useful
CC for epidemiological studies
XX
SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 1489;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTCGGCTTGAAC 31
DB 391 CTCACGGCATCGTCAGTTCGGCTTGAAC 421

```

```

RESULT 3
AAZ36113
ID AAZ36113 standard; DNA; 31 BP.
XX
AC AAZ36113;
XX
DT 11-FEB-2000 (first entry)
XX
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
OS Synthetic.
OS Escherichia coli.
XX
PN WO9555908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
PA (SNFI ) PASTEUR SANOPI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
WPI; 2000-013443/01.
XX
New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 5; Page 27; 48pp; French.
XX
AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of
CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
XX
SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 87.1%; Score 27; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTCGGCTTGG 27
DB 5 CTCACGGCATCGTCAGTTCGGCTTGG 31

RESULT 4
AAZ36110/c
ID AAZ36110 standard; DNA; 26 BP.
XX
AC AAZ36110;
XX
DT 11-FEB-2000 (first entry)
XX
DE Primer derived from a nucleic acid sequence specific to EHEC.

```

```

XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
XX Synthetic.
OS Escherichia coli.
OS
XX WO9955908-A2.
XX
XX PD 04-NOV-1999.
XX
XX PF 27-APR-1999; 99WO-FR001000.
XX
XX PR 28-APR-1998; 98FR-00005329.
XX
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX PI Frechon DTM, Laure FC, Thierry D;
XX
XX DR WPI; 2000-013443/01.
XX
XX PT New nucleic acid containing sequences specific to enterohemorrhagic
XX Escherichia coli, particularly serotype O157:H7, used for detecting these
XX bacteria in food.
XX
XX PS Claim 5; Page 27; 48pp; French.
XX
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
XX specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
XX derived from two sequences. The first (AAZ36101) is 99.9% homologous to
XX the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
XX 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
XX The second sequence (AAZ36102) is associated with the presence of
XX virulence factors: enterohemolysine (ehly) and intimin (eae). Nucleotides
XX 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
XX for virulence proteins of Shigella flexneri. Both sequences are of
XX plasmid origin. The fragments are used as PCR primers and probes for the
XX detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
XX in human or animal samples, foods or the environment. The fragments are
XX also useful for epidemiological studies
XX
XX SQ Sequence 26 BP; 8 A; 7 C; 8 G; 3 T; 0 U; 0 Other;

Query Match 67.7%; Score 21; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 TCGTCAGTTGGCGCTTGGAC 31
Db 26 TCGTCAGTTGGCGCTTGGAC 6

RESULT 5
ADB68745/C
ID ADB68745 standard; DNA; 664 BP.
XX
XX AC ADB68745;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE Pseudomonas chlororaphis phzI DNA.
XX
XX KW quorum sensing; lux homologue; luxI; ds.
XX
XX OS Pseudomonas chlororaphis.
XX
XX PN WO2003057902-A2.
XX
XX PD 17-JUL-2003.
XX
XX PF 08-JAN-2003; 2003WO-US000479.
XX
XX
XX 08-JAN-2002; 2002US-0346531P.
XX 07-JAN-2003; 2003US-00338110.
XX (FRAU ) FRAUNHOFER USA INC.
XX
XX PI Fuhrmann JJ, Romesser JA;
XX
XX XX WPI; 2003-618102/58.
XX
XX DR Detecting quorum sensing potential of a Gram-negative bacterium in a
XX sample comprises performing a polymerase chain reaction using nucleic
XX acids extracted from a sample containing a microorganism.
XX
XX PS Disclosure; Fig 10; 86pp; English.
XX
XX CC The invention relates to a novel method for detecting the quorum sensing
XX potential of a microorganism in a sample which comprises performing PCR
XX using nucleic acids extracted from a sample containing at least one type
XX of microorganism. The method may be useful for detecting the quorum
XX sensing potential of a microorganism in a sample by amplifying a fragment
XX of a lux gene or homologue. The current sequence is that of the luxI
XX homologue DNA of the invention.
XX
XX SQ Sequence 664 BP; 162 A; 219 C; 167 G; 116 T; 0 U; 0 Other;

Query Match 65.2%; Score 20.2; DB 10; Length 664;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTTGCGGCTTG 27
Db 421 CAACGGCATCGTCAGTTGCGGCTTG 397

RESULT 6
AAZ36111/C
ID AAZ36111 standard; DNA; 26 BP.
XX
XX AC AAZ36111;
XX
XX DT 11-FEB-2000 (first entry)
XX
XX DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
XX PCR primer; probe; ss.
XX
XX OS Synthetic.
XX OS Escherichia coli.
XX
XX PN WO9955908-A2.
XX
XX PD 04-NOV-1999.
XX
XX PF 27-APR-1999; 99WO-FR001000.
XX
XX PR 28-APR-1998; 98FR-00005329.
XX
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX PI Frechon DTM, Laure FC, Thierry D;
XX
XX DR WPI; 2000-013443/01.
XX
XX PT New nucleic acid containing sequences specific to enterohemorrhagic
XX Escherichia coli, particularly serotype O157:H7, used for detecting these
XX bacteria in food.
XX
XX PS Claim 5; Page 27; 48pp; French.
XX
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
XX specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are

```

CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katp gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of
CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virk gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
XX
SQ Sequence 26 BP; 7 A; 7 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 64.5%; Score 20; DB 3; Length 26;

Best Local Similarity 100.0%; Pred. No. 94; Mismatches 0; Indels 0; Gaps 0;

QY 12 CGTCAGTTGCGGCTTGGAC 31

DB 26 CGTCAGTTGCGGCTTGGAC 7

RESULT 7

AAL49962

ID AAL49962 standard; DNA; 636 BP.

XX

AC AAL49962;

XX

XX 10-DEC-2002 (first entry)

XX

XX Human B1153 expression in allergic disease related human DNA #4.

XX

XX Human; allergy; B1153; differential expression; antiallergic; asthma;
KW antiasthmatic; antiinflammatory; atopic skin inflammation; ds.

XX

XX Homo sapiens.

XX

XX WO200250269-A1.

XX

XX 27-JUN-2002.

XX

XX 21-DEC-2001; 2001WO-JP011286.

XX

XX 21-DEC-2000; 2000JP-00389476.

XX

XX (GENO-) GENOX RES INC.

XX

XX (NIGE-) JAPAN GEN NAT CHILDREN'S HOSPITAL.

XX

XX Matsumoto Y, Imai Y, Oshida T, Sugita Y, Nagasu T, Tsujimoto G;

XX

XX WPI; 2002-713252/77.

XX

XX Examination of allergic diseases comprises detecting gene B1153 over-
PT expressed in T cells of allergy patients for diagnosis treatment and
PT investigation of atopic skin inflammation and asthma.

XX

XX Example 7; Page 89-90; 102pp; Japanese.

XX

XX The present invention relates to a method of examining allergic diseases
CC which comprises comparing the expression level of gene B1153 in allergy
CC patients with the expression level in healthy subjects. The method is
CC useful for the treatment, prevention, diagnosis and study of allergic
CC diseases including atopic skin inflammation and asthma. The present
CC sequence is a human sequence described in the exemplification of the
CC invention

XX

SQ Sequence 636 BP; 221 A; 104 C; 132 G; 179 T; 0 U; 0 Other;

Query Match

Best Local Similarity 64.5%; Score 20; DB 6; Length 636;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCAACGGCATCGTCAGTTGCGGCTTGGGA 29
DB 341 TCAAAGGAAACGTCAGTTGCTGATTGGA 368

RESULT 8

AAL49970

ID AAL49970 standard; DNA; 824 BP.

XX

AC AAL49970;

XX

XX 10-DEC-2002 (first entry)

XX

XX Human B1153 related coding sequence.

XX

XX Human; allergy; B1153; differential expression; antiallergic; asthma;

KW antiasthmatic; antiinflammatory; atopic skin inflammation; gene; ds.

XX

XX Homo sapiens.

XX

XX Key Location/Qualifiers

XX

XX 50..824

XX

XX /*tag= a

XX

XX /partial

XX

XX /note= "no stop codon"

XX

XX WO200250269-A1.

XX

XX 27-JUN-2002.

XX

XX 21-DEC-2001; 2001WO-JP011286.

XX

XX 21-DEC-2000; 2000JP-00389476.

XX

XX (GENO-) GENOX RES INC.

XX

XX (NIGE-) JAPAN GEN NAT CHILDREN'S HOSPITAL.

XX

XX Matsumoto Y, Imai Y, Oshida T, Sugita Y, Nagasu T, Tsujimoto G;

XX

XX WPI; 2002-713252/77.

XX

XX P-PSDB; AAO19419.

XX

XX Examination of allergic diseases comprises detecting gene B1153 over-
PT expressed in T cells of allergy patients for diagnosis treatment and
PT investigation of atopic skin inflammation and asthma.

XX

XX Example 7; Page 95-97; 102pp; Japanese.

XX

XX The present invention relates to a method of examining allergic diseases
CC which comprises comparing the expression level of gene B1153 in allergy
CC patients with the expression level in healthy subjects. The method is
CC useful for the treatment, prevention, diagnosis and study of allergic
CC diseases including atopic skin inflammation and asthma. The present
CC sequence is a human coding sequence described in the exemplification of
CC the invention

XX

SQ Sequence 824 BP; 275 A; 148 C; 179 G; 222 T; 0 U; 0 Other;

Query Match

Best Local Similarity 64.5%; Score 20; DB 6; Length 824;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;


```
DT 05-FEB-2003 (first entry)
XX Glyceraldehyde-3-phosphate dehydrogenase 11.99 cDNA.
DE
DE
XX
XX Glyceraldehyde-3-phosphate dehydrogenase 11.99; nervous disease;
KW metabolism disturbance disease; development disturbance; tumour;
KW inflammation; immunological disease; gene; ss.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FH 104..433
FT /tag= a
FT /product= "Glyceraldehyde-3-phosphate dehydrogenase
FT 11.99"
XX
XX CN1358847-A.
XX
XX 17-JUL-2002.
XX
XX 13-DEC-2000; 2000CN-00127876.
XX
XX 13-DEC-2000; 2000CN-00127876.
XX
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2002-733568/80.
XX P-PSDB; ABP57886.
XX
XX Novel polypeptide-glyceraldehyde-3-phosphate dehydrogenase 11.99 and
XX polynucleotide for encoding said polypeptide.
XX
XX Claim 6; Page 25-26 (Disclosure); 32pp; Chinese.
XX
XX The invention relates to a novel polypeptide, glyceraldehyde-3-phosphate
XX dehydrogenase 11.99, and the polynucleotide encoding it. The polypeptide
XX is useful for treating several diseases, such as metabolism disturbance
XX disease, nervous disease, development disturbance, tumours, inflammations
XX and immunological disease. The invention also discloses an antagonist for
XX resisting the polypeptide and its therapeutic action, and the application
XX of the polynucleotide. The present sequence encodes the glyceraldehyde-3-
XX phosphate dehydrogenase 11.99 of the invention
XX
XX Sequence 1018 BP; 328 A; 198 C; 222 G; 270 T; 0 U; 0 Other;
SQ
Query Match 64.5%; Score 20; DB 6; Length 1018;
Best Local Similarity 82.1%; Pred. No. 1.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2 TCAACGGCATCGTCAGTTGCGGCTTGA 29
Db 584 TCAAGGAAACGTCAGTTGCTGATTGA 611
RESULT 10
ADB80402
ID ADB80402 standard; DNA; 1143 BP.
XX
XX ADB80402;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX Human MDDT gene SEQ ID NO: 89.
DE
XX human; molecule for disease detection and treatment; MDDT; cardiant;
KW cycostatic; neuroprotective; gene therapy; cardiovascular disorder;
KW neurological disorder; cancer; gene; ds.
XX
XX Homo sapiens.
OS
XX WO2003016497-A2.
PN
```

```
XX 27-FEB-2003.
PD
XX 16-AUG-2002; 2002WO-US026411.
PF
XX 17-AUG-2001; 2001US-03131115P.
PR 31-AUG-2001; 2001US-0316785P.
PR 17-SEP-2001; 2001US-0323407P.
PR 28-SEP-2001; 2001US-0326022P.
PR 05-APR-2002; 2002US-0370707P.
PR 10-MAY-2002; 2002US-0379862P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Baughn MR, Yao MG, Swarnakar A, Ison CH, Walia NK, Tang YT;
XX Becha SD, Yue H, Lee EA, Gorvad AE, Burford N, Forsythe IJ;
XX Emmerling BM, Griffin JA, Hafalia AJA, Barroso I, Elliott VS, Lu DAM;
XX Duggan BM, Warren BA, Xu Y, Lee SY, Richardson TW, Thangavelu K;
XX Lee S, Lehr-Mason PM, Kearney L, Gandhi AR, Zebartjadian Y;
XX Au-Young JK, Tran UK, Lindquist EA, Gietzen KJ;
XX
XX WPI; 2003-268321/26.
XX P-PSDB; ADB80353.
XX
XX New molecules for disease detection and treatment polypeptide, useful for
XX preparing a composition for diagnosing or treating e.g. cardiovascular or
XX neurological disorders.
XX
XX Claim 5; Page 336-337; 345pp; English.
XX
XX The invention relates to a novel isolated polypeptide associated with
XX molecules for disease detection and treatment (MDDT). A polypeptide of
XX the invention has cardiant, cytostatic, and neuroprotective activity, and
XX may have a use in gene therapy, and as a MDDT-Antagonist or MDDT-Agonist.
XX The polypeptide is useful for preparing a composition for diagnosing or
XX treating a disease or condition associated with decreased expression or
XX overexpression of functional MDDT e.g. cardiovascular or neurological
XX disorders or cancer. The sequences shown in ADB80363-ADB80411 represent
XX MDDT genes of the invention.
XX
XX Sequence 1143 BP; 381 A; 212 C; 229 G; 321 T; 0 U; 0 Other;
SQ
Query Match 64.5%; Score 20; DB 9; Length 1143;
Best Local Similarity 82.1%; Pred. No. 1.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2 TCAACGGCATCGTCAGTTGCGGCTTGA 29
Db 553 TCAAGGAAACGTCAGTTGCTGATTGA 580
RESULT 11
ADL03792/c
ID ADL03792 standard; DNA; 1770 BP.
XX
XX ADL03792;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX DNA encoding a M. catarrhalis protein #1478.
DE
XX ds; gene; Moraxella catarrhalis; infection.
XX
XX Moraxella catarrhalis.
OS
XX US6673910-B1.
XX
XX 06-JAN-2004.
PD
XX 04-APR-2000; 2000US-00540236.
XX
XX 08-APR-1999; 99US-0128416P.
XX
```

```
PA (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL;
XX WPI; 2004-178127/17.
DR P-PSDB; ADL05712.
XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
XX caused by Moraxella catarrhalis.
XX Disclosure; SEQ ID NO 1478; 429pp; English.
XX The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
CC catarrhalis protein.
XX Sequence 1770 BP; 494 A; 400 C; 417 G; 459 T; 0 U; 0 Other;
SQ Query Match 64.5%; Score 20; DB 12; Length 1770;
Best Local Similarity 82.1%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTCAACGGCATCGTCAGTGGCGTTGG 28
Db ||||| ||||| ||||| ||||| |||||
648 CTCAACGGCATATCGGTTTCGGTTGG 621

RESULT 12
ADA53102
ID ADA53102 standard; cDNA; 1858 BP.
XX
AC ADA53102;
XX
XX 20-NOV-2003 (first entry)
XX Human coding sequence, SEQ ID 670.
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX Homo sapiens.
XX EP1293569-A2.
XX 19-MAR-2003.
XX 21-MAR-2002; 2002EP-00006586.
XX 14-SEP-2001; 2001JP-00328381.
XX 24-JAN-2002; 2002US-0350435P.
XX (HELJ-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-395539/38.
DR P-PSDB; ADA54741.
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX Claim 1; SEQ ID NO 670; 205pp; English.
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC
```

```
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX Sequence 1858 BP; 612 A; 333 C; 404 G; 509 T; 0 U; 0 Other;
SQ Query Match 64.5%; Score 20; DB 10; Length 1858;
Best Local Similarity 82.1%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TCAACGGCATCGTCAGTGGCGTTGGA 29
Db ||||| ||||| ||||| ||||| |||||
572 TCAACGGAAACGTCAGTTGCTGATTGGA 599

RESULT 13
AAL49947
ID AAL49947 standard; DNA; 3596 BP.
XX
AC AAL49947;
XX
XX 10-DEC-2002 (first entry)
XX Human B1153 coding sequence.
XX Human; allergy; B1153; differential expression; antiallergic; asthma;
KW antiaethmatic; antiinflammatory; atopic skin inflammation; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 86..2980
CDS /*tag= a
FT /product= "B1153"
FT
XX WO200250269-A1.
XX 27-JUN-2002.
XX 21-DEC-2001; 2001WO-JP011286.
XX 21-DEC-2000; 2000JP-00389476.
XX (GENO-) GENOX RES INC.
XX (NIGE-) JAPAN GEN NAT CHILDREN'S HOSPITAL.
XX Matsumoto Y, Imai Y, Oshida T, Sugita Y, Nagasu T, Tsujimoto G;
PI WPI; 2002-713252/77.
DR P-PSDB; AAO19418.
XX Examination of allergic diseases comprises detecting gene B1153 over-
PT expressed in T cells of allergy patients for diagnosis treatment and
PT investigation of atopic skin inflammation and asthma.
XX Claim 1; Page 64-76; 102pp; Japanese.
XX The present invention relates to a method of examining allergic diseases
CC which comprises comparing the expression level of gene B1153 in allergy
CC patients with the expression level in healthy subjects. The method is
CC useful for the treatment, prevention, diagnosis and study of allergic
CC diseases including atopic skin inflammation and asthma. The present
CC sequence is the human B1153 coding sequence
XX Sequence 3596 BP; 1206 A; 619 C; 742 G; 1029 T; 0 U; 0 Other;
SQ Query Match 64.5%; Score 20; DB 6; Length 3596;
Best Local Similarity 82.1%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TCAACGGCATCGTCAGTGGCGTTGGA 29
Db ||||| ||||| ||||| ||||| |||||
565 TCAAGGAAACGTCAGTTGCTGATTGGA 592
```

RESULT 14
ADQ64982
ID ADQ64982 standard; cDNA; 3614 BP.
XX
XX
AC ADQ64982;
XX
XX
DT 07-OCT-2004 (first entry)
XX
XX
DE Novel human cDNA sequence #2143.
XX
XX
KW ss: Gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytosolic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
XX
OS Homo sapiens.
XX
XX
PN EP1440981-A2.
XX
XX
PD 28-JUL-2004.
XX
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
XX
PR 21-JAN-2003; 2003JP-00102206.
XX
XX
PR 09-MAY-2003; 2003JP-00131392.
XX
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX
PI Isogai T, Sugiyama T, Otauki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
XX
DR WPI; 2004-535376/52.
XX
XX
DR P-PSDB; ADQ67170.
XX
XX
SQ Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX
PS Claim 1; SEQ ID NO 2143; 2449pp; English.
XX
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 3614 BP; 1189 A; 631 C; 752 G; 1042 T; 0 U; 0 Other;
Query Match 64.5%; Score 20; DB 12; Length 3614;
Best Local Similarity 82.1%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 2 TCACAGGCATCGTCAGTTCGGCTTGA 29
DB 610 TCAAAGGAACGTCAGTTCGCTGATTGA 637
RESULT 15
AAF28530/c
ID AAF28530 standard; DNA; 23210 BP.
XX
XX
AC AAF28530;
XX
XX
DT 04-APR-2001 (first entry)
XX
XX
DE Genomic fragment #17.
XX
XX
KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;

KW bronchopulmonary; endocarditis; meningitis; ss.
XX
XX
OS Moraxella catarrhalis.
XX
XX
PN WO200078968-A2.
XX
XX
PD 28-DEC-2000.
XX
XX
PF 16-JUN-2000; 2000WO-US016649.
XX
XX
PR 18-JUN-1999; 99US-0140121P.
XX
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX
PI Lagace RE, Patterson C, Berg KL;
XX
XX
DR WPI; 2001-041427/05.
XX
XX
PT Genomic library for identifying diagnostic and therapeutic compositions,
PT and for identifying virulence factors, regulatory elements and drug
PT targets, comprises Moraxella catarrhalis nucleic acids.
XX
XX
PS Claim 1; Page 141-146; 545pp; English.
XX
XX
CC The present invention relates to a Moraxella catarrhalis genomic library
CC comprising of a combination of 41 nucleic acid molecules (see AAF28514-
CC AAF28554). The library has a number of uses described in the
CC specification e.g. is useful for identifying diagnostic and therapeutic
CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC aerobic, gram-negative diplococcus, normally found among the bacterial
CC flora of human upper airways. M. catarrhalis is known to cause acute,
CC localised infections such as otitis media, sinusitis and bronchopulmonary
CC infection and life-threatening, systemic diseases including endocarditis
CC and meningitis
XX
SQ Sequence 23210 BP; 7108 A; 4482 C; 4951 G; 6668 T; 0 U; 1 Other;
Query Match 64.5%; Score 20; DB 4; Length 23210;
Best Local Similarity 82.1%; Pred. No. 2.4e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 CTCACGGCATCGTCAGTTCGGCTTGG 28
DB 16504 CTCACGGCATATCGTTCGGTTTGG 16477
RESULT 16
ADS63488
ID ADS63488 standard; cDNA; 3237 BP.
XX
XX
AC ADS63488;
XX
XX
DT 02-DEC-2004 (first entry)
XX
XX
DE Bacterial polynucleotide #15475.
XX
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
XX
OS Bacteria.
XX
XX
PN US2003233675-A1.
XX
XX
PD 18-DEC-2003.
XX
XX
PF 20-FEB-2003; 2003US-00369493.
XX
XX
PR 21-FEB-2002; 2002US-0360039P.

```
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX XX WPI; 2004-061375/06.
XX XX
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plants with
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
XX CC providing improved plant growth and development under at least one stress
XX CC condition. Improved lignin production or improved galactomannan
XX CC production. This sequence represents a bacterial polynucleotide used in
XX CC the scope of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification but was obtained in electronic
XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 3237 BP; 600 A; 1034 C; 1083 G; 520 T; 0 U; 0 Other;

Query Match 63.9%; Score 19.8; DB 13; Length 3237;
Best Local Similarity 77.4%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CTCACGGCATCGTCAGTTGGCGCTTGGAC 31
Db 562 CTGATCGACATCGTCGGCCGCGCGCTGGAC 592

RESULT 17
ADSG4237
ID ADS64237 standard; cDNA; 3240 BP.
XX AC ADS64237;
XX XX
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polynucleotide #16224.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polynucleotide; gene; ss.
XX OS Bacteria.
XX XX

PN US2003233675-A1.
XX PD
XX PF 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX XX
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX XX WPI; 2004-061375/06.
XX XX
XX CC New recombinant DNA construct comprising a promoter positioned to provide
XX CC for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 39911; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plants with
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
XX CC providing improved plant growth and development under at least one stress
XX CC condition. Improved lignin production or improved galactomannan
XX CC production. This sequence represents a bacterial polynucleotide used in
XX CC the scope of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification but was obtained in electronic
XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 3240 BP; 601 A; 1035 C; 1084 G; 520 T; 0 U; 0 Other;

Query Match 63.9%; Score 19.8; DB 13; Length 3240;
Best Local Similarity 77.4%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CTCACGGCATCGTCAGTTGGCGCTTGGAC 31
Db 562 CTGATCGACATCGTCGGCCGCGCGCTGGAC 592

RESULT 18
ADSG3858
ID ADS63858 standard; cDNA; 3240 BP.
XX AC ADS63858;
XX XX
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polynucleotide #15845.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
```

KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX Bacteria.
 XX US2003233675-A1.
 XX 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX New recombinant DNA construct comprising a promoter positioned to provide
 XX for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 39532; 122pp; English.
 XX The invention relates to a recombinant DNA construct comprising a
 XX promoter functional in a plant cell, where the promoter is positioned to
 XX provide for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source. The invention also relates to a transformed plant
 XX comprising the recombinant DNA construct and a method of producing a
 XX transformed plant having an improved property. The plant is a crop plant
 XX such as maize or soybean. The method of producing a transformed plant
 XX having an improved property comprises transforming a plant with the
 XX recombinant DNA construct and growing the transformed plant, where the
 XX polynucleotide or polypeptide is useful for improving plant properties.
 XX The recombinant DNA construct is useful for producing plants with
 XX improved plant properties, e.g. improved cold, heat or drought tolerance,
 XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 XX increased resistance to plant disease, better growth rate by modification
 XX of the cell cycle pathway with plant growth regulators, increased rate of
 XX homologous recombination, modified seed oil or protein yield and/or
 XX content, improved yield by modification of carbohydrate, nitrogen or
 XX phosphorus use and/or uptake, by modification of photosynthesis or by
 XX providing improved plant growth and development under at least one stress
 XX condition, improved lignin production or improved galactomannan
 XX production. This sequence represents a bacterial polynucleotide used in
 XX the scope of the invention. Note: The sequence data for this patent did
 XX not form part of the printed specification but was obtained in electronic
 XX format from USPTO at seqdata.uspto.gov/sequence.html.
 XX Sequence 3240 BP; 601 A; 1035 C; 1084 G; 520 T; 0 U; 0 Other;
 XX Query Match 63.9%; Score 19.8; DB 13; Length 3240;
 XX Best Local Similarity 77.4%; Pred. No. 2.2e+02;
 XX Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 CTCACGGCATCGTCAGTTCGGCTGGAAC 31
 |||||
 Db 562 CTGATCGACATCGTCGGCGCGGCTGGAAC 592
 |||||
 RESULT 19
 ACH83759
 ID ACH83759 standard; DNA; 242 BP.
 XX
 AC ACH83759;
 XX

DT 29-JUN-2004 (first entry)
 XX Human genome derived single exon probe #16954.
 XX Human; probe; ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX Homo sapiens.
 OS US2003194704-A1.
 XX 16-OCT-2003.
 XX 03-APR-2002; 2002US-00029386.
 XX 03-APR-2002; 2002US-00029386.
 XX (PENN/) PENN S G.
 XX (RANK/) RANK D R.
 XX (HANZ/) HANZEL D K.
 XX Penn SG, Rank DR, Hanzel DK;
 XX WPI; 2004-119264/12.
 XX New human genome-derived single exon nucleic acid probes useful for human
 XX gene expression analysis, for identifying or characterizing alternative
 XX splicing events, for assessing genomic alterations or as tools for
 XX surveying tissues.
 XX Claim 1; SEQ ID NO 16954; 80pp; English.
 XX The invention relates to a nucleic acid probe for measuring human gene
 XX expression, comprising any of the 27,400 fully defined nucleotide
 XX sequences in the specification, or their complements or fragments, and
 XX encoding at least 8 amino acids of any of the 6888 amino acid sequences
 XX fully defined in the specification. The probe is a single exon probe that
 XX hybridises under high stringency conditions to a nucleic acid molecule
 XX expressed in human cells or tissues. Also included are a spatially-
 XX addressable set of single exon nucleic acid probes for measuring human
 XX gene expression (comprising a plurality of single exon nucleic acid
 XX probes cited above, where each of the plurality of probes is separately
 XX and addressably isolatable or amplifiable from the plurality), a single
 XX exon microarray for measuring human gene expression, a method of
 XX measuring human gene expression, a vector comprising the single exon
 XX probe cited above, an ORF-encoded peptide comprising at least 8
 XX contiguous amino acids of any of the above-mentioned amino acid
 XX sequences (optionally with conservative amino acid substitutions), an
 XX isolated antibody that binds specifically to a peptide cited above,
 XX methods of selling and/or licensing single exon probes or microarrays to
 XX a customer desiring to measure gene expression, a method of providing
 XX human gene expression data by subscription, and a computer-readable
 XX storage medium which contains a database having a plurality of records
 XX (each record including data on the expression of a single exon probe
 XX cited above). The probe, methods and apparatus are useful in gene
 XX expression analysis. The probes may be used as tools for surveying
 XX tissues to detect the presence of expressed messages that contain their
 XX specific exon, or in constructing genome-derived single exon microarrays.
 XX In addition, the probes are used in identifying and characterising
 XX alternative splicing events, in detecting and characterising gross
 XX alterations in the genomic locus that includes their exon, in assessing
 XX smaller genomic alterations, in priming the synthesis of nucleic acid,
 XX or in expressing the ORF-encoded peptide. The present sequence is a human
 XX single exon probe of the invention. Note: The sequence data for this
 XX patent did not form part of the printed specification, but was obtained
 XX in electronic format directly from USPTO at
 XX seqdata.uspto.gov/sequence.html?docID=20030194704
 XX Sequence 242 BP; 39 A; 87 C; 75 G; 41 T; 0 U; 0 Other;
 XX Query Match 62.6%; Score 19.4; DB 12; Length 242;
 XX Best Local Similarity 79.3%; Pred. No. 2.3e+02;
 XX Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTGGCGCTTGGAAC 31
Db 116 CTACGGCGTGGTCAGCTCGCGCTTGGACC 144

RESULT 20

ACH70059
ID ACH70059 standard; DNA; 509 BP.

XX AC ACH70059;

XX DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon probe #3254.

XX KW Human; probe; ss; gene expression; single exon probe; microarray;

XX KW alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX PN US2003194704-A1.

XX PD 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

XX PA (RANK/) RANK D R.

XX PA (HANZ/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX PT New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.

XX PS Claim 15; SEQ ID NO 3254; 80pp; English.

XX CC The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above). The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing

CC smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

SQ Sequence 509 BP; 94 A; 163 C; 163 G; 89 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 12; Length 509;

Best Local Similarity 79.3%; Pred. No. 2.6e+02;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTGGCGCTTGGAAC 31

Db 149 CTACGGCGTGGTCAGCTCGCGCTTGGACC 177

RESULT 21

AAC34842/C

ID AAC34842 standard; DNA; 1391 BP.

XX AC AAC34842;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 8084.

XX KW Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 04-MAY-1999; 99US-0132407P.

XX PR 05-MAY-1999; 99US-0132484P.

XX PR 06-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132487P.

XX PR 11-MAY-1999; 99US-0132863P.

XX PR 14-MAY-1999; 99US-0134256P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 14-MAY-1999; 99US-0134370P.

XX PR 18-MAY-1999; 99US-0134768P.

XX PR 19-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135533P.

XX PR 24-MAY-1999; 99US-0135629P.

XX PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139889P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140911P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145184P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147132P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.

```
PR 29-OCT-1999; 99US-0162142P.
Query Match 62.6%; Score 19.4; DB 3; Length 1391;
Best Local Similarity 79.3%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTCAACGGCATCGTCAGTTCGGCTTGA 29
    ||||| ||||| ||||| ||||| |||||
Db 683 CTCGAAGGCATTGTCAGTTACAGCTTTGA 655

RESULT 22
AAD18767
ID AAD18767 standard; cDNA; 2700 BP.
AC AAD18767;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human ANIC-BP-1 ligand encoding cDNA.
XX
KW Human; acute neuronal induced calcium binding protein type 1 ligand;
KW ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
KW Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 363..2435
FT /*tag= a
FT /product= "Human ANIC-BP-1 ligand"
XX
PN WO200170771-A2.
XX
PD 27-SEP-2001.
XX
PF 20-MAR-2001; 2001WO-EP003149.
XX
PR 21-MAR-2000; 2000EP-00106110.
XX
PA (MERE ) MERCK PATENT GMBH.
XX
PI Den Daas I, Duecker K, Hock B;
XX
DR WPI; 2001-607519/69.
XX
DR P-PSDB; AAE10857.
XX
PT Novel acute neuronal induced calcium binding protein type 1 ligand
PT polypeptides, useful in the treatment of stroke, head trauma, multiple
PT sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
PT injury.
XX
PS Claim 4; Page 36-39; 46pp; English.
XX
CC The invention relates to human acute neuronal induced calcium binding
CC protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.
CC Sequences of the invention are useful for treating human diseases
CC including stroke, head trauma, multiple sclerosis, Parkinson's disease,
CC Alzheimer's disease and spinal cord injury. They are also useful as
CC vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
CC soluble receptors. Polynucleotides of the invention are useful as
CC diagnostic reagents, for chromosome localization studies, and as valuable
CC tools for tissue expression studies. They are also useful in gene
CC therapy. The present sequence is human ANIC-BP-1 ligand cDNA
XX
SQ Sequence 2700 BP; 436 A; 901 C; 857 G; 506 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 4; Length 2700;
Best Local Similarity 79.3%; Pred. No. 3.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CAACGGCATCGTCAGTTCGGCTTGAAC 31

PR 29-OCT-1999; 99US-0162142P.
Query Match 62.6%; Score 19.4; DB 3; Length 1391;
Best Local Similarity 79.3%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTCAACGGCATCGTCAGTTCGGCTTGA 29
    ||||| ||||| ||||| ||||| |||||
Db 683 CTCGAAGGCATTGTCAGTTACAGCTTTGA 655

RESULT 23
ADR58978
ID ADR58978 standard; cDNA; 3017 BP.
XX
AC ADR58978;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human Elk1 phosphorylation/Elk1 kinase activation cDNA - SEQ ID 81.
XX
KW Elk1 phosphorylation; Elk1 kinase activation; antiinflammatory;
KW immunosuppressive; cytosolic; anti-HIV; neuroprotective; osteopathic;
KW dermatological; inflammation; autoimmune; cancer; infectious; bone; AIDS;
KW neurodegenerative; ischaemic injury; graft-versus-host disease; skin;
KW IGA nephritis; purpuric nephritis; proliferative nephritis;
KW fulminant hepatitis; gene therapy; human; ss; gene.
XX
OS Homo sapiens.
XX
PN WO2004072277-A2.
XX
PD 26-AUG-2004.
XX
PF 13-FEB-2004; 2004WO-JP001587.
XX
PR 13-FEB-2003; 2003JP-00034875.
PR 14-FEB-2003; 2003US-0447320P.
XX
PA (ASAH-) ASahi Kasei Pharma Corp.
XX
PI Matsuzaki O, Matsuda A;
XX
DR WPI; 2004-642303/62.
DR P-PSDB; ADR58979.
XX
PT New Elk1 phosphorylation related protein, useful for diagnosing,
PT treating, or preventing a disease, e.g. inflammation, autoimmune
PT diseases, cancers, AIDS, neurodegenerative diseases, or bone and skin
PT diseases.
XX
PS Claim 4; SEQ ID NO 81; 415pp; English.
XX
CC The invention relates to a novel polypeptide that acts to phosphorylate
CC Elk1 and/or activate a kinase which phosphorylates Elk1. The polypeptide
CC of the invention demonstrates antiinflammatory, immunosuppressive,
CC cytosolic, anti-HIV, neuroprotective, osteopathic and dermatological
CC activities and may be useful for diagnosing, treating or preventing a
CC disease, such as inflammation, autoimmune disease, cancer, infectious
CC disease, bone disease, AIDS, neurodegenerative disease, ischaemic injury,
CC graft-versus-host disease, skin disease, IGA nephritis, purpuric
CC nephritis, proliferative nephritis or fulminant hepatitis. Furthermore,
CC the polypeptide and polynucleotide of the invention may be utilised
CC during gene therapy. The current sequence is that of a human Elk1
CC phosphorylation/Elk1 kinase activation-related cDNA of the invention.
XX
SQ Sequence 3017 BP; 562 A; 929 C; 947 G; 579 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 13; Length 3017;
Best Local Similarity 79.3%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CAACGGCATCGTCAGTTCGGCTTGAAC 31
    ||||| ||||| ||||| ||||| |||||
Db 1478 CTACGGCTGTCAGTCGGCTGGACC 1506

RESULT 24
ADR58980
ID ADR58980 standard; cDNA; 6862 BP.
```


AC ADR59890;
XX 18-NOV-2004 (first entry)
XX Human Elkl phosphorylation/Elkl kinase activation cDNA - SEQ ID 83.
DE Elkl phosphorylation; Elkl kinase activation; antiinflammatory;
XX immunosuppressive; cytostatic; anti-HIV; neuroprotective; osteopathic;
KW dermatological; inflammation; autoimmune; cancer; infectious; bone; AIDS;
KW neurodegenerative; ischaemic injury; graft-versus-host disease; skin;
KW IGA nephritis; purpuric nephritis; proliferative nephritis;
KW fulminant hepatitis; gene therapy; human; ss; gene.
XX Homo sapiens.
XX WO2004072277-A2.
XX 26-AUG-2004.
XX 13-FEB-2004; 2004WO-JP001587.
XX 13-FEB-2003; 2003JP-00034875.
XX 14-FEB-2003; 2003US-0447320P.
XX (ASAH-) ASahi KASEI PHARMA CORP.
XX Matsuzaki O, Matsuda A;
PI WPI; 2004-642303/62.
DR P-PSDB; ADR59891.
XX New Elkl phosphorylation related protein, useful for diagnosing,
PT treating, or preventing a disease, e.g. inflammation, autoimmune
PT diseases, cancers, AIDS, neurodegenerative diseases, or bone and skin
XX diseases.
XX Claim 4; SEQ ID NO 83; 415pp; English.
XX The invention relates to a novel polypeptide that acts to phosphorylate
CC Elkl and/or activate a kinase which phosphorylates Elkl. The polypeptide
CC of the invention demonstrates antiinflammatory, immunosuppressive,
CC cytostatic, anti-HIV, neuroprotective, osteopathic and dermatological
CC activities and may be useful for diagnosing, treating or preventing a
CC disease, such as inflammation, autoimmune disease, cancer, infectious
CC disease, bone disease, AIDS, neurodegenerative disease, ischaemic injury,
CC graft-versus-host disease, skin disease, IGA nephritis, purpuric
CC nephritis, proliferative nephritis or fulminant hepatitis. Furthermore,
CC the polypeptide and polynucleotide of the invention may be utilised
CC during gene therapy. The current sequence is that of a human Elkl
CC phosphorylation/Elkl kinase activation-related cDNA of the invention.
XX SQ Sequence 6862 BP; 1442 A; 1999 C; 1953 G; 1468 T; 0 U; 0 Other;
Query Match 62.6%; Score 19.4; DB 13; Length 6862;
Best Local Similarity 79.3%; Pred. No. 3.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 CAACGGCATCGTCAGTTGCGGCTTGGAAAC 31
DB 1432 CTACGGCTGTCAGTCGCGGCTTGGACC 1460
RESULT 25
ADR59004
ID ADR59004 standard; cDNA; 6908 BP.
XX
AC ADR59004;
XX 18-NOV-2004 (first entry)
XX Human Elkl phosphorylation/Elkl kinase activation cDNA - SEQ ID 107.
KW Elkl phosphorylation; Elkl kinase activation; antiinflammatory;
KW immunosuppressive; cytostatic; anti-HIV; neuroprotective; osteopathic;
KW dermatological; inflammation; autoimmune; cancer; infectious; bone; AIDS;
KW neurodegenerative; ischaemic injury; graft-versus-host disease; skin;
KW IGA nephritis; purpuric nephritis; proliferative nephritis;
KW fulminant hepatitis; gene therapy; human; ss; gene.
XX Homo sapiens.
XX WO2004072277-A2.
XX 26-AUG-2004.
XX 13-FEB-2004; 2004WO-JP001587.
XX 13-FEB-2003; 2003JP-00034875.
XX 14-FEB-2003; 2003US-0447320P.
XX (ASAH-) ASahi KASEI PHARMA CORP.
XX Matsuzaki O, Matsuda A;
PI WPI; 2004-642303/62.
DR P-PSDB; ADR59005.
XX New Elkl phosphorylation related protein, useful for diagnosing,
PT treating, or preventing a disease, e.g. inflammation, autoimmune
PT diseases, cancers, AIDS, neurodegenerative diseases, or bone and skin
XX diseases.
XX Claim 4; SEQ ID NO 107; 415pp; English.
XX The invention relates to a novel polypeptide that acts to phosphorylate
CC Elkl and/or activate a kinase which phosphorylates Elkl. The polypeptide
CC of the invention demonstrates antiinflammatory, immunosuppressive,
CC cytostatic, anti-HIV, neuroprotective, osteopathic and dermatological
CC activities and may be useful for diagnosing, treating or preventing a
CC disease, such as inflammation, autoimmune disease, cancer, infectious
CC disease, bone disease, AIDS, neurodegenerative disease, ischaemic injury,
CC graft-versus-host disease, skin disease, IGA nephritis, purpuric
CC nephritis, proliferative nephritis or fulminant hepatitis. Furthermore,
CC the polypeptide and polynucleotide of the invention may be utilised
CC during gene therapy. The current sequence is that of a human Elkl
CC phosphorylation/Elkl kinase activation-related cDNA of the invention.
XX SQ Sequence 6908 BP; 1446 A; 2018 C; 1968 G; 1476 T; 0 U; 0 Other;
Query Match 62.6%; Score 19.4; DB 13; Length 6908;
Best Local Similarity 79.3%; Pred. No. 3.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 CAACGGCATCGTCAGTTGCGGCTTGGAAAC 31
DB 1478 CTACGGCTGTCAGTCGCGGCTTGGACC 1506
Search completed: March 11, 2005, 04:19:56
Job time : 48.3508 secs

This Page Blank (uspto)


```
; ORGANISM: Candida albicans
US-09-248-796A-7463

Query Match      65.2%; Score 20.2; DB 4; Length 645;
Best Local Similarity 88.0%; Pred. No. 16;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TCAACGGCATCGTCAGTTCGGGCTT 26
Db 151 TCAACGGCATCGACAGTTCGCTT 175

RESULT 3
US-09-540-236-1478/c
; Sequence 1478, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709-2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1478
; LENGTH: 1770
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-1478

Query Match      64.5%; Score 20; DB 4; Length 1770;
Best Local Similarity 82.1%; Pred. No. 24;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCAACGGCATCGTCAGTTCGGGCTTGG 28
Db 648 CTCAACGGCATATCGGTTTCGGTTTGG 621

RESULT 4
US-09-596-002-17/c
; Sequence 17, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 23210
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 28
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 17
; PUBLICATION INFORMATION:
US-09-596-002-17

Query Match      64.5%; Score 20; DB 4; Length 23210;
Best Local Similarity 82.1%; Pred. No. 43;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCAACGGCATCGTCAGTTCGGGCTTGG 28
```

```
Db 16504 CTCAACGGCATATCGGTTTCGGTTTGG 16477

RESULT 5
US-08-872-855-3/c
; Sequence 3, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,855
; FILING DATE: 11-JUN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MAA-003.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2055 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-872-855-3

Query Match      61.3%; Score 19; DB 3; Length 2055;
Best Local Similarity 81.5%; Pred. No. 71;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 ACGGCATCGTCAGTTCGGGCTTGGAAC 31
Db 1287 ACGGCACGGCACATCGGCTTGGACC 1261

RESULT 6
US-09-641-612-10/c
; Sequence 10, Application US/09641612
; Patent No. 6703221
; GENERAL INFORMATION:
; APPLICANT: Vivien Chan et al.
; TITLE OF INVENTION: NOTCH RECEPTOR LIGANDS AND USES THEREOF
; FILE REFERENCE: PPO-1602.002 / 200130.498
; CURRENT APPLICATION NUMBER: US/09/641,612
; CURRENT FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 2055
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-641-612-10
```

Query Match 61.3%; Score 19; DB 4; Length 2055;
Best Local Similarity 81.5%; Pred. No. 71;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 ACGGCATCGTCAGTTGGCGCTTGAAC 31
||||| ||||| ||||| ||||| |||||
Db 1287 ACGGCAGCGGCACATCGCGCTTGACC 1261

RESULT 7

US-08-872-855-1/c
; Sequence 1, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
; NUMBER OF SEQUENCES: 23.
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872,855
; FILING DATE: 11-JUN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: NAA-003.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 338..2392

US-08-872-855-1

Query Match 61.3%; Score 19; DB 3; Length 2800;
Best Local Similarity 81.5%; Pred. No. 76;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 ACGGCATCGTCAGTTGGCGCTTGAAC 31
||||| ||||| ||||| ||||| |||||
Db 1624 ACGGCAGCGGCACATCGCGCTTGACC 1598

RESULT 8

US-09-423-753-4/c
; Sequence 4, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753
; CURRENT FILING DATE: 1999-12-30

; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3339
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (277)..(2331)
; NAME/KEY: sig_peptide
; LOCATION: (277)..(354)
; NAME/KEY: mat_peptide
; LOCATION: (355)..(2331)
US-09-423-753-4

Query Match 61.3%; Score 19; DB 4; Length 3339;
Best Local Similarity 81.5%; Pred. No. 79;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 ACGGCATCGTCAGTTGGCGCTTGAAC 31
||||| ||||| ||||| ||||| |||||
Db 1563 ACGGCAGCGGCACATCGCGCTTGACC 1537

RESULT 9

US-08-545-562A-4
; Sequence 4, Application US/08545562A
; Patent No. 5840529
; GENERAL INFORMATION:
; APPLICANT: SEIDAH, Nabil G.
; APPLICANT: DAY, Robert
; APPLICANT: CHRETIEN, Michel
; TITLE OF INVENTION: MAMMALIAN PRO-HORMONE CONVERTASE
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,562A
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/510,347
; FILING DATE: 02-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/517,015
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 200702.90025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```
; MOLECULE TYPE: cDNA
US-08-545-562A-4

Query Match      60.6%; Score 18.8; DB 2; Length 285;
Best Local Similarity 76.7%; Pred. No. 56;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TCACGGCATCGTCAGTTGGCGCTTGGAC 31
Db 107 TCAATGACATCTACAGTTCAGCTGGGAC 136

RESULT 10
US-09-949-016-18801
; Sequence 18801, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18801
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18801

Query Match      59.4%; Score 18.4; DB 4; Length 601;
Best Local Similarity 78.6%; Pred. No. 1e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 AACGGCATCGTCAGTTGGCGCTTGGAC 31
Db 479 AACGGAGCGTCAGTATCATCCTGGGAC 506

RESULT 11
US-09-949-016-66642
; Sequence 66642, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66642
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-66642

Query Match      59.4%; Score 18.4; DB 4; Length 601;
Best Local Similarity 78.6%; Pred. No. 1e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 AACGGCATCGTCAGTTGGCGCTTGGAC 31
Db 479 AACGGAGCGTCAGTATCATCCTGGGAC 506

RESULT 12
US-09-902-540-4174
; Sequence 4174, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4174
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4174

Query Match      59.4%; Score 18.4; DB 4; Length 1011;
Best Local Similarity 78.6%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTGGCGCTTGG 28
Db 769 CTCATCGGCATGCTCAGCGCGGCTTCG 796

RESULT 13
US-09-782-587B-2
; Sequence 2, Application US/09782587B
; Patent No. 6806063
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, ANDERS H.
; APPLICANT: ANDERSON, KIM V.
; APPLICANT: BORNAES, CLAUS
; TITLE OF INVENTION: FACTOR VII OR VIIA-LIKE MOLECULES
; FILE REFERENCE: 31-001100US
; CURRENT APPLICATION NUMBER: US/09/782,587B
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: PA 2000 00218
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 60/184,036
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/241,916
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)..(1332)
US-09-782-587B-2

Query Match      59.4%; Score 18.4; DB 4; Length 1338;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTGGCGCTTGG 28
Db 769 CTCATCGGCATGCTCAGCGCGGCTTCG 796
```

Db 1186 CTCACGGGATCGTCAGTCGGGCCAGG 1213

RESULT 14

US-09-782-587B-4
; Sequence 4, Application US/09782587B
; Patent No. 6806063
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, ANDERS H.
; APPLICANT: ANDERSON, KIM V.
; APPLICANT: BORNAES, CLAUD
; TITLE OF INVENTION: FACTOR VII OR VIIA-LIKE MOLECULES
; FILE REFERENCE: 31-001100US
; CURRENT APPLICATION NUMBER: US/09/782,587B
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: PA 2000 00218
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 60/184,036
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/241,916
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1357
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Expression
; OTHER INFORMATION: cassette for expression of FVII in mammalian cells
US-09-782-587B-4

Query Match 59.4%; Score 18.4; DB 4; Length 1357;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTCACGGGATCGTCAGTCGGGCTTGG 28

Db 1199 CTCACGGGATCGTCAGTCGGGCCAGG 1226

RESULT 15

US-09-902-540-9314
; Sequence 9314, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9314
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9314

Query Match 59.4%; Score 18.4; DB 4; Length 1572;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTCACGGGATCGTCAGTCGGGCTTGG 28

Db 721 CGCACGGGATCGTCAGTCGGGCGGTGG 748

RESULT 16

US-09-902-540-995/c

; Sequence 995, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 995

; LENGTH: 11992

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-995

Query Match 59.4%; Score 18.4; DB 4; Length 11992;

Best Local Similarity 78.6%; Pred. No. 2.e+02;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTCACGGGATCGTCAGTCGGGCTTGG 28

Db 3579 CGCACGGGATCGTCAGTCGGGCGGTGG 3552

RESULT 17

US-09-902-540-1208

; Sequence 1208, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 1208

; LENGTH: 22301

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-1208

Query Match 59.4%; Score 18.4; DB 4; Length 22301;

Best Local Similarity 78.6%; Pred. No. 2.3e+02;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTCACGGGATCGTCAGTCGGGCTTGG 28

Db 7679 CTCATCGGATCGTCAGCGGGGCTTGG 7706

RESULT 18

US-09-949-016-13702/c

; Sequence 13702, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13702
; LENGTH: 54251
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(54251)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13702

Query Match      59.4%; Score 18.4; DB 4; Length 54251;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      4 AACGGCATCGTCAGTTGCGGCTTGGAAAC 31
Db      48347 AACGGCAGCGTCAGTATCACCTGGGAAC 48320

RESULT 19
US-09-949-016-11807/c
; Sequence 11807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11807
; LENGTH: 54252
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(54252)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11807

Query Match      59.4%; Score 18.4; DB 4; Length 54252;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      4 AACGGCATCGTCAGTTGCGGCTTGGAAAC 31
Db      48348 AACGGCAGCGTCAGTATCACCTGGGAAC 48321

RESULT 20
US-09-489-039A-1737/c
; Sequence 1737, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
```

```
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1737
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1737

Query Match      58.1%; Score 18; DB 4; Length 774;
Best Local Similarity 80.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      6 CGGCATCGTCAGTTGCGGCTTGGAAAC 31
Db      681 CGGCATCGTCAGTTCCCATTTGGAAAC 656

RESULT 21
US-09-543-681A-2390/c
; Sequence 2390, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2390
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2390

Query Match      58.1%; Score 18; DB 4; Length 1071;
Best Local Similarity 80.8%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 AACGGCATCGTCAGTTGCGGCTTGGGA 29
Db      795 AACGGCATCTTCAGGTGGAGATTGGA 770

RESULT 22
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
```


US-09-103-840A-2

Query Match 58.1%; Score 18; DB 3; Length 4403765;
Best Local Similarity 80.8%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 CGCATCGTCAGTTCGGCTTGGAAAC 31
DB 604042 CGGCATCGGAGTTCGGCTCGGGCAC 604067

RESULT 23

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS
; FILE REFERENCES: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match 58.1%; Score 18; DB 3; Length 4411529;
Best Local Similarity 80.8%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 CGCATCGTCAGTTCGGCTTGGAAAC 31
DB 602682 CGGCATCGGAGTTCGGCTCGGGCAC 602707

RESULT 24

US-09-103-840A-2
; Patent No. 5218099
; APPLICANT: REYES, GREGORY R.;BRADLEY, DANIEL W.;RABIN,LINDA;
; FRY, KIRK
; TITLE OF INVENTION: POST-TRANSFUSION, NON-A, NON-B HEPATITIS
; VIRUS POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/372,711
; FILING DATE: 28-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 350,570; 334,701; 228,334; 215,728; 846,757
; FILING DATE: 11-MAY-1989
; APPLICATION NUMBER: 334,701
; FILING DATE: 06-APR-1989
; APPLICATION NUMBER: 228,334
; FILING DATE: 04-AUG-1988
; APPLICATION NUMBER: 215,728
; FILING DATE: 06-JUL-1988
; APPLICATION NUMBER: 846,757
; FILING DATE: 01-APR-1986
; SEQ ID NO: 6
; LENGTH: 669

5218099-6

Query Match 57.4%; Score 17.8; DB 6; Length 669;
Best Local Similarity 90.5%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 GCATCGTCAGTTCGGCTTGG 28
DB 443 GCATCGTCAGTTCGCTTGG 423

RESULT 25
5218099-6/c
; Patent No. 5218099
; APPLICANT: REYES, GREGORY R.;BRADLEY, DANIEL W.;RABIN,LINDA;
; FRY, KIRK
; TITLE OF INVENTION: POST-TRANSFUSION, NON-A, NON-B HEPATITIS
; VIRUS POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/372,711
; FILING DATE: 28-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 350,570; 334,701; 228,334; 215,728; 846,757
; FILING DATE: 11-MAY-1989
; APPLICATION NUMBER: 334,701
; FILING DATE: 06-APR-1989
; APPLICATION NUMBER: 228,334
; FILING DATE: 04-AUG-1988
; APPLICATION NUMBER: 215,728
; FILING DATE: 06-JUL-1988
; APPLICATION NUMBER: 846,757
; FILING DATE: 01-APR-1986
; SEQ ID NO: 6
; LENGTH: 669

5218099-6

Query Match 57.4%; Score 17.8; DB 6; Length 669;
Best Local Similarity 90.5%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 GCATCGTCAGTTCGGCTTGG 28
DB 443 GCATCGTCAGTTCGCTTGG 423

Search completed: March 11, 2005, 13:13:05
Job time : 22.557 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 154.869 Seconds

(without alignments)
1190.710 Million cell updates/sec

Title: US-09-674-277-12

Perfect score: 31

Sequence: 1 ctcaacgcagtcagtcagtcggttggaac 31

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20.2	65.2	664	17	US-10-338-110-22
2	20	64.5	571	18	US-10-425-115-58469
3	20	64.5	636	18	US-10-239-734-17
4	20	64.5	824	18	US-10-239-734-25
5	20	64.5	1858	17	US-10-094-749-670
6	20	64.5	3596	18	US-10-239-734-1
7	20	64.5	23210	17	US-10-672-787-17
8	19.8	63.9	3237	17	US-10-369-493-39162
9	19.8	63.9	3240	17	US-10-369-493-39532
10	19.8	63.9	3240	17	US-10-369-493-39911
11	19.4	62.6	242	16	US-10-029-386-16954

12	19.4	62.6	509	16	US-10-029-386-3254	Sequence 3254, Ap
13	19.4	62.6	584	18	US-10-425-115-114661	Sequence 114661, Ap
14	19.4	62.6	2700	15	US-10-239-079-1	Sequence 1, Appl
C 15	19.2	61.9	1156	17	US-10-282-122A-36496	Sequence 36496, A
16	19.2	61.9	1247	17	US-10-425-114-35365	Sequence 35365, A
17	19.2	61.9	1252	18	US-10-425-115-86277	Sequence 86277, A
C 18	19.2	61.9	1278	17	US-10-374-780A-1323	Sequence 1323, Ap
C 19	19.2	61.9	1278	17	US-10-414-699B-1472	Sequence 1472, Ap
C 20	19.2	61.9	1362	17	US-10-369-493-44549	Sequence 44549, A
C 21	19.2	61.9	1562	17	US-10-282-122A-38840	Sequence 38840, A
C 22	19.2	61.9	1686	17	US-10-282-122A-39500	Sequence 39500, A
C 23	19	61.3	434	17	US-10-424-599-6242	Sequence 6242, Ap
C 24	19	61.3	724	9	US-09-764-898-45	Sequence 45, Appl
C 25	19	61.3	835	18	US-10-437-963-55487	Sequence 55487, A
C 26	19	61.3	1584	15	US-10-225-630-3	Sequence 3, Appl
C 27	19	61.3	1632	17	US-10-369-493-35920	Sequence 35920, A
C 28	19	61.3	2055	16	US-10-417-719-3	Sequence 3, Appl
C 29	19	61.3	2055	18	US-10-644-548-10	Sequence 10, Appl
C 30	19	61.3	2058	18	US-10-731-741-10	Sequence 10, Appl
C 31	19	61.3	2159	14	US-10-028-072-87	Sequence 87, Appl
C 32	19	61.3	2159	14	US-10-140-808-87	Sequence 87, Appl
C 33	19	61.3	2159	14	US-10-121-049-87	Sequence 87, Appl
C 34	19	61.3	2159	14	US-10-123-904-87	Sequence 87, Appl
C 35	19	61.3	2159	14	US-10-140-470-87	Sequence 87, Appl
C 36	19	61.3	2159	14	US-10-175-746-87	Sequence 87, Appl
C 37	19	61.3	2159	14	US-10-176-918-87	Sequence 87, Appl
C 38	19	61.3	2159	14	US-10-176-921-87	Sequence 87, Appl
C 39	19	61.3	2159	14	US-10-127-884-213	Sequence 213, App
C 40	19	61.3	2159	14	US-10-137-865-87	Sequence 87, Appl
C 41	19	61.3	2159	14	US-10-140-474-87	Sequence 87, Appl
C 42	19	61.3	2159	14	US-10-142-431-87	Sequence 87, Appl
C 43	19	61.3	2159	14	US-10-143-114-87	Sequence 87, Appl
C 44	19	61.3	2159	14	US-10-230-163-213	Sequence 213, App
C 45	19	61.3	2159	14	US-10-230-338-213	Sequence 213, App

ALIGNMENTS

RESULT 1

US-10-338-110-22/c
; Sequence 22, Application US/10338110
; Publication No. US20040023254A1
; GENERAL INFORMATION:
; APPLICANT: Fuhrmann, Jeffrey J.
; APPLICANT: Ramesser, James A.
; TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial
; FILE REFERENCE: HER-0056
; CURRENT APPLICATION NUMBER: US/10/338,110
; CURRENT FILING DATE: 2003-01-07
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Pseudomonas aureofaciens
US-10-338-110-22

Query Match 65.2%; Score 20.2; DB 17; Length 664;
Best Local Similarity 88.0%; Pred. No. 52;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CAACGCGCATCGTCAGTTCGGCTTG 27

Db 421 CAACGCGCATCGTCAGTTCGGCTTG 397

RESULT 2

US-10-425-115-58469
; Sequence 58469, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihui
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 58469
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_153319C.1
US-10-425-115-58469

Query Match 64.5%; Score 20; DB 18; Length 571;
Best Local Similarity 82.1%; Pred. No. 63;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCAACGGCATCGTCAGTTCGGCTTGGG 29
||||| ||||| ||||| ||||| |||||
Db 518 TCAACGCACATCGTCAACGGCGTCTTGGG 545

RESULT 3
US-10-239-734-17
; Sequence 17, Application US/10239734
; Publication No. US20040161746A1
; GENERAL INFORMATION:
; APPLICANT: GENOX RESEARCH, INC.
; APPLICANT: JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF AGENCY OF NATIONAL CENTER FOR
; APPLICANT: CHILD HEALTH AND DEVELOPMENT
; APPLICANT: Tsujimoto, Yoshiko
; APPLICANT: Matsumoto, Gozoh
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Sugita, Yuji
; APPLICANT: Oshida, Tadahiho
; APPLICANT: Imai, Yukiko
; TITLE OF INVENTION: Method of Testing For Allergic Disease
; FILE REFERENCE: SHIMIZU-07379
; CURRENT APPLICATION NUMBER: US/10/239,734
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/JPO1/11286
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 2000-389476 JP
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-734-17

Query Match 64.5%; Score 20; DB 18; Length 636;
Best Local Similarity 82.1%; Pred. No. 63;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCAACGGCATCGTCAGTTCGGCTTGGG 29
||||| ||||| ||||| ||||| |||||
Db 341 TCAAGGAAACGTCAGTTCGCTGATTGGA 368

RESULT 4
US-10-239-734-25
; Sequence 25, Application US/10239734
; Publication No. US20040161746A1
; GENERAL INFORMATION:
; APPLICANT: GENOX RESEARCH, INC.
; APPLICANT: JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF AGENCY OF NATIONAL CENTER FOR

; APPLICANT: CHILD HEALTH AND DEVELOPMENT
; APPLICANT: Matsumoto, Yoshiko
; APPLICANT: Tsujimoto, Gozoh
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Sugita, Yuji
; APPLICANT: Oshida, Tadahiho
; APPLICANT: Imai, Yukiko
; TITLE OF INVENTION: Method of Testing For Allergic Disease
; FILE REFERENCE: SHIMIZU-07379
; CURRENT APPLICATION NUMBER: US/10/239,734
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/JPO1/11286
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 2000-389476 JP
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(823)
; OTHER INFORMATION:
US-10-239-734-25

Query Match 64.5%; Score 20; DB 18; Length 824;
Best Local Similarity 82.1%; Pred. No. 65;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TCAACGGCATCGTCAGTTCGGCTTGGG 29
||||| ||||| ||||| ||||| |||||
Db 529 TCAAGGAAACGTCAGTTCGCTGATTGGA 556

RESULT 5
US-10-094-749-670
; Sequence 670, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 670
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-094-749-670

Query Match 64.5%; Score 20; DB 17; Length 1858;
Best Local Similarity 82.1%; Pred. No. 70;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TCAACGGCATCGTCAGTTGGCGCTTGA 29
|||||
Db 572 TCAAGGAACGTCAGTTGCTGATTGA 599
|||||

RESULT 6

US-10-239-734-1
; Sequence 1, Application US/10239734
; Publication No. US20040161746A1
; GENERAL INFORMATION:
; APPLICANT: GENOX RESEARCH, INC.
; APPLICANT: JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF AGENCY OF NATIONAL CENTER FOR
; APPLICANT: CHILD HEALTH AND DEVELOPMENT
; APPLICANT: Matsumoto, Yoshiko
; APPLICANT: Tsujimoto, Gozon
; APPLICANT: Nagasu, Takeshi
; APPLICANT: Sugita, Yuji
; APPLICANT: Oshida, Tadahiro
; APPLICANT: Imai, Yukio
; TITLE OF INVENTION: Method of Testing For Allergic Disease
; FILE REFERENCE: SHIMIZU-07379
; CURRENT APPLICATION NUMBER: US/10/239,734
; CURRENT FILING DATE: 2002-09-24
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/JP01/11286
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 2000-389476 JP
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)..(2980)
; OTHER INFORMATION:
US-10-239-734-1

Query Match 64.5%; Score 20; DB 18; Length 3596;
Best Local Similarity 82.1%; Pred. No. 74;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TCAACGGCATCGTCAGTTGGCGCTTGA 29
|||||
Db 565 TCAAGGAACGTCAGTTGCTGATTGA 592
|||||

RESULT 7

US-10-672-787-17/c
; Sequence 17, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELITRA.025C1
; CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 09/596,002
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 23210
; TYPE: DNA

; ORGANISM: Moraxella catarrhalis

; FEATURE:
; NAME/KEY: unsure
; LOCATION: 28
; OTHER INFORMATION: a or g or c or t, unknown, or other
US-10-672-787-17

Query Match 64.5%; Score 20; DB 17; Length 23210;
Best Local Similarity 82.1%; Pred. No. 88;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTCAACGGCATCGTCAGTTGGCGCTTGG 28
|||||
Db 16504 CTCAACGGCATATCGGTTTCGGTTGG 16477
|||||

RESULT 8

US-10-369-493-39162
; Sequence 39162, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 39162
; LENGTH: 3237
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
US-10-369-493-39162

Query Match 63.9%; Score 19.8; DB 17; Length 3237;
Best Local Similarity 77.4%; Pred. No. 90;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CTCAACGGCATCGTCAGTTGGCGCTTGAAC 31
|||||
Db 562 CTGATCGACATCGTCGCGCGCGCTGGAAC 592
|||||

RESULT 9

US-10-369-493-39532
; Sequence 39532, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 39532
; LENGTH: 3240
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
US-10-369-493-39532

Query Match 63.9%; Score 19.8; DB 17; Length 3240;
Best Local Similarity 77.4%; Pred. No. 90;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTCAAGGCATCGTCAGTTGCGGCTTGGAAAC 31
|||
DB 562 CTGATCGACATCGTCGGCCGCGCCTGGAAAC 592

RESULT 10
US-10-369-493-39911
; Sequence 39911, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 39911
; LENGTH: 3240
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
US-10-369-493-39911

Query Match 63.9%; Score 19.8; DB 17; Length 3240;
Best Local Similarity 77.4%; Pred. No. 90;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CTCAAGGCATCGTCAGTTGCGGCTTGGAAAC 31
|||
DB 562 CTGATCGACATCGTCGGCCGCGCCTGGAAAC 592

RESULT 11
US-10-029-386-16954
; Sequence 16954, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Rank, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16954
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR17.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.6
; OTHER INFORMATION: SWISSPROT HIT: O27606, EVALUE 4.90e+00
; OTHER INFORMATION: EST HUMAN HIT: BE260736.1, EVALUE 1.00e-123
; OTHER INFORMATION: NT HIT: g114774017, EVALUE 0.00e+00
US-10-029-386-16954

Query Match 62.6%; Score 19.4; DB 16; Length 242;
Best Local Similarity 79.3%; Pred. No. 11e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CAACGGCATCGTCAGTTGCGGCTTGGAAAC 31
|||
DB 116 CTACGGCCTGGTCAGCTGCGGCTGGACC 144

RESULT 12
US-10-029-386-3254
; Sequence 3254, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3254
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR17.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.6
; OTHER INFORMATION: SWISSPROT HIT: Q01844, EVALUE 3.60e+00
; OTHER INFORMATION: EST HUMAN HIT: BE260736.1, EVALUE 1.00e-122
; OTHER INFORMATION: NT HIT: g114774013, EVALUE 0.00e+00
US-10-029-386-3254

Query Match 62.6%; Score 19.4; DB 16; Length 509;
Best Local Similarity 79.3%; Pred. No. 11e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CAACGGCATCGTCAGTTGCGGCTTGGAAAC 31
|||
DB 149 CTACGGCCTGGTCAGCTGCGGCTGGACC 177

RESULT 13
US-10-425-115-114661
; Sequence 114661, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 114661
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_36052C.1
; OTHER INFORMATION: NT HIT: g114774017, EVALUE 0.00e+00
US-10-425-115-114661

Query Match 62.6%; Score 19.4; DB 18; Length 584;

Best Local Similarity 79.3%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTTGGCGCTTGGAAAC 31
|||||
Db 371 CAACAGCATCGTCAGTCGCGGTGAGGACC 399
|||||

RESULT 14

US-10-239-079-1

; Sequence 1, Application US/10239079
; Publication No. US2003014846A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP1-1ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (363)..(2432)
; FEATURE:
; OTHER INFORMATION: Description of ANIC-BP-1 protein ligand

US-10-239-079-1

Query Match 62.6%; Score 19.4; DB 15; Length 2700;
Best Local Similarity 79.3%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTTGGCGCTTGGAAAC 31
|||||
Db 1766 CTACGGCTGTCAGTCGCGGCTTGGACC 1794
|||||

RESULT 15

US-10-282-122A-36496/c

; Sequence 36496, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36496
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (151)..(152)
; OTHER INFORMATION: n=g, a, t or c
US-10-282-122A-36496

Query Match 61.9%; Score 19.2; DB 17; Length 1156;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTCACGGCATCGTCAGTTGGCGC 24
|||||
Db 293 CTCACGGCAGCATCAGTTGGCGC 270
|||||

RESULT 16

US-10-425-114-35365

; Sequence 35365, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 35365
; LENGTH: 1247
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73014D08_FLI

US-10-425-114-35365

Query Match 61.9%; Score 19.2; DB 17; Length 1247;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TCACGGCATCGTCAGTTGGCGCT 25
|||||
Db 689 TCACGGCGTCGTCAGTTGGCGCT 712
|||||

RESULT 17

US-10-425-115-86277

; Sequence 86277, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 86277
; LENGTH: 1252
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_178695C.1
; US-10-425-115-86277

Query Match      61.9%; Score 19.2; DB 18; Length 1252;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 TCAACGGCATCGTCAGTTGCGGCT 25
      |||||||
Db      691 TCCACGGCGTCGTCAGTTGCGGCT 714

RESULT 18
US-10-374-780A-1323/c
; Sequence 1323, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riemann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1323
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G987
;

```

```

US-10-374-780A-1323
Query Match      61.9%; Score 19.2; DB 17; Length 1278;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 TCAACGGCATCGTCAGTTGCGGCT 25
      |||||||
Db      577 TCCACGGCGTCGTCAGTTGCGGCT 554

RESULT 19
US-10-412-699B-1472/c
; Sequence 1472, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riemann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumamoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1472
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-412-699B-1472
Query Match      61.9%; Score 19.2; DB 17; Length 1278;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 TCAACGGCATCGTCAGTTGCGGCT 25

```



```
Query Match          61.9%; Score 19.2; DB 17; Length 1686;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTCAAGCGCATCGTCAGTTGCGGC 24
Db 785 CTCACGGCAGCATCAGTTGCGGC 762

RESULT 23
US-10-424-599-6242
; Sequence 6242, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 6242
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105646C.1
US-10-424-599-6242

Query Match          61.3%; Score 19; DB 17; Length 434;
Best Local Similarity 81.5%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCAAGCGCATCGTCAGTTGCGGCTTG 27
Db 238 CTCAAGGCTTCCTCAGTCACGGCTTG 264

RESULT 24
US-09-764-898-45/c
; Sequence 45, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-898-45

Query Match          61.3%; Score 19; DB 9; Length 724;
Best Local Similarity 81.5%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 ACGGCATCGTCAGTTGCGGCTTGGAAC 31
Db 557 ACGCAGCGGCACATCGCGGCTTGACC 531

RESULT 25
US-10-437-963-55487/c
; Sequence 55487, Application US/10437963
; Publication No. US20040123343A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 55487
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57490C.1
US-10-437-963-55487

Query Match          61.3%; Score 19; DB 18; Length 835;
Best Local Similarity 81.5%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CAACGGCATCGTCAGTTGCGGCTTGA 29
Db 667 CGACGGCACCCTCAGCTACGGCATGGA 641

Search completed: March 12, 2005, 00:25:16
Job time : 155.869 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 410.213 Seconds
(without alignments)
2876.537 Million cell updates/sec

Title: US-09-674-277-12
Perfect score: 31
Sequence: 1 ctcaaggcatcgcagtcggttggaac 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21.4	69.0	312	4	BM401635 JLLG04F S
2	21	67.7	385	7	CK391871 K0836G10-
3	21	67.7	527	5	BW582626 BW582626
4	21	67.7	532	4	BI745607 r182h12.y
5	21	67.7	633	5	BW580869 BW580869
6	21	67.7	656	2	BB093704 BB093704
7	21	67.7	749	7	CR769342 CR769342
8	21	67.7	847	4	BI110178 BI110178
9	21	67.7	975	5	BUS16442 AGENCOURT
10	21	67.7	1653	3	AK034825 Mus muscu
C 11	20.6	66.5	516	7	CK744057 ltu01-14m
C 12	20.4	65.8	493	9	CL155476 104 341-1
C 13	20.4	65.8	510	4	BM401667 BM401667
14	20.4	65.8	595	4	BM344014 BM344014
15	20.4	65.8	609	6	CA077416 SCQSAM103
C 16	20.4	65.8	641	8	BZ341420 ic45c05.g
17	20.4	65.8	816	6	CA148723 SCJLRZ101
C 18	20.4	65.8	934	7	CF885756 trico033xh
C 19	20.4	65.8	1151	9	AL441494 T3 end of
20	20.2	65.2	271	8	AZ214463 Sheared D
21	20.2	65.2	414	8	BH254127 SALK 0160
C 22	20.2	65.2	548	9	TA176A11Q
23	20	64.5	337	9	BX214909 Danio rer
24	20	64.5	345	1	AA985360 am80h08.s

25	20	64.5	381	1	AI207748
26	20	64.5	469	7	CO295193
27	20	64.5	625	7	CN367932
28	20	64.5	629	9	BX242611
C 29	20	64.5	653	8	AZ566286
30	20	64.5	669	4	BG530421
31	20	64.5	681	7	CF874806
C 32	20	64.5	734	6	CA300647
33	20	64.5	757	6	CB904069
34	20	64.5	765	6	CB988288
35	20	64.5	768	6	CB985373
36	20	64.5	797	7	CK603475
37	20	64.5	878	5	BU166514
38	20	64.5	883	4	BI458483
39	20	64.5	1000	5	BX384762
40	20	64.5	1089	4	BG704735
41	20	64.5	1166	4	BM543044
42	20	64.5	3411	3	CR597934
43	19.8	63.9	440	1	AU070098
C 44	19.8	63.9	532	4	BM031605
45	19.8	63.9	563	7	CN951072

ALIGNMENTS

RESULT 1 BM401635/c 312 bp mRNA linear EST 01-MAY-2002
LOCUS JLLG04F Snake Bothrops insularis library IL3 Bothrops insularis
DEFINITION cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401635
VERSION BM401635.1 GI:20376263
KEYWORDS EST.
SOURCE Bothrops insularis (Island Jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 312)
AUTHORS Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PubMed 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL01A (see Reference)
Seq primer: M13F.
Location/Qualifiers
1. .312
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/notes="Organ: venom glands; Vector: pGEM11zf+; Site 1: Eco RI; Site 2: Not I; Sug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adaptors and directionally cloned in pGEM11zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or

cluster was obtained through Blast searches (e-value < e-05)."

```

ORIGIN
Query Match          69.0%; Score 21.4; DB 4; Length 312;
Best Local Similarity 80.6%; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTCCGCGCTTGGAC 31
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 CTCAGTCGATCTTCAGATTAGGCTTGGAC 253

RESULT 2
CK391871
LOCUS CK391871 385 bp mRNA linear EST 29-DEC-2003
DEFINITION K0836G10-5 NIA Mouse 8.5-dpc Whole Embryo cDNA Library (Long) Mus
            musculus cDNA clone NIA:K0836G10 IMAGE:30081873 5', mRNA sequence.
ACCESSION CK391871.1 GI:40382330
VERSION EST.
KEYWORDS Mus musculus
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 385)
Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
PUBLISHED
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgaun.grc.nia.nih.gov
Plate: K0836 row: G column: 10
Seq primer: M13 Reverse
High quality sequence stop: 385
POLYA=No.

FEATURES
source Location/Qualifiers
1..385
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="niaEST:K0836G10-5"
/clones="NIA:K0836G10 IMAGE:30081873"
/tissue_type="whole embryo including extraembryonic
tissues at 8.5-days postcoitum"
/dev_stage="8.5-days postcoitum"
/lab_hosts="DH10B"
/clone_lib="NIA Mouse 8.5-dpc Whole Embryo cDNA Library
(Long)"
/note="vector: pSPOR1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgaun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
extracted from a pool of 13 embryos at 8.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen:
5'-TGACTAGTCTCAGATCGGAGCGCCGCTTTT-3'] from
9.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker Lu-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon

```

100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPOR1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

```

ORIGIN
Query Match          67.7%; Score 21; DB 7; Length 385;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTCACGGCATCGTCAGTTCCGCGCTTGGAC 29
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 189 CTGAAGGCATCCTCAGTTCCGCGCTTGGAC 217

RESULT 3
BW582626
LOCUS BW582626 527 bp mRNA linear EST 01-SEP-2004
DEFINITION BW582626 Yutaka Satou unpublished cDNA library (csef1) Ciona
            savignyi cDNA clone csef021e07 3', mRNA sequence.
ACCESSION BW582626
VERSION EST.
KEYWORDS Ciona savignyi
SOURCE Ciona savignyi
ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
1 (bases 1 to 527)
Yamada,L., Satoh,N. and Satou,Y.
Expressed genes in Ciona savignyi (Yamada, Satoh, Satou)
Unpublished (2004)
Contact: Yutaka Satou
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113
Email: yutaka@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source Location/Qualifiers
1..527
/organism="Ciona savignyi"
/mol_type="mRNA"
/db_xref="taxon:51511"
/clone="csef021e07"
/dev_stage="egg"
/clone_lib="Yutaka Satou unpublished cDNA library (csef1)"

ORIGIN
Query Match          67.7%; Score 21; DB 5; Length 527;
Best Local Similarity 82.8%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CAACGGCATCGTCAGTTCCGCGCTTGGAC 31
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 CATCGGCATCGTCATTTCGGTTGGTAC 142

RESULT 4
BI745607
LOCUS BI745607 532 bp mRNA linear EST 25-SEP-2001
DEFINITION BI745607 rK82h12.y3 Meloidogyne javanica egg pAMP1 v6 Chiapelli McCarter
            Meloidogyne javanica cDNA 5', similar to TR:Q17446 Q17446 SIMILARITY
            TO CDC2/CDC28 SUBFAMILY OF SER/THR PROTEIN KINASES. [1] ;, mRNA
            sequence.
ACCESSION BI745607
VERSION BI745607.1 GI:15767409
KEYWORDS EST.
SOURCE Meloidogyne javanica (root-knot nematode)
ORGANISM Meloidogyne javanica
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.

```

REFERENCE
AUTHORS

1 (bases 1 to 532)
McCarte, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
Wylie, T., Dater, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R., and Wilson, R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarte JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Brandi Chiapelli and Dr. James
McCarte (bchiapel@watson.wustl.edu & jmcarte@watson.wustl.edu) at
Washington University, St. Louis. DNA Sequencing by: Washington
University Genome Sequencing Center St. Louis.
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 396.
Location/Qualifiers
1. 532
/organism="Meloiodogyne javanica"
/mol_type="mRNA"
/db_xref="taxon:6303"
/dev_stage="enriched for eggs"
/lab_host="DH10B"
/clone_lib="Meloiodogyne javanica egg pAMP1 v6 Chiapelli
McCarte"
/note="Vector: pAMP1 (Gibco); The library was constructed
by Brandi Chiapelli and Dr. James McCarte at Washington
University, St. Louis. The cDNA was made by using Dynabead
oligo-dT priming (Dynal). PCR based library using a
modified protocol from the SMART PCR cDNA Synthesis Kit
from Clontech. Directionally cloned into the UDG sites of
pAMP1. Nematodes were provided by Dr. David Bird of North
Carolina State University."

ORIGIN

Query Match 67.7%; Score 21; DB 4; Length 532;
Best Local Similarity 82.8%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TCAACGGCATCGTCAGTTGCGGCTTGGAA 30
|||||
Db 361 TCAACAGCTTCGGGAGTTGCAGCTTGGAA 389
|||||

RESULT 5
BW580869
LOCUS
DEFINITION
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BW580869 Yutaka Satou unpublished cDNA library (csefl) Ciona
savignyi cDNA clone csef01lc15 3', mRNA sequence.
BW580869.1 GI:51761303
EST.
Ciona savignyi
Ciona savignyi
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
1 (bases 1 to 633)
Yamada, L., Satoh, N. and Satou, Y.
Expressed genes in Ciona savignyi (Yamada, Satoh, Satou)
Unpublished (2004)
Contact: Yutaka Satou
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 633)
Yamada, L., Satoh, N. and Satou, Y.
Expressed genes in Ciona savignyi (Yamada, Satoh, Satou)
Unpublished (2004)
Contact: Yutaka Satou
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113

FEATURES
source

Location/Qualifiers
1. 633
/organism="Ciona savignyi"
/mol_type="mRNA"
/db_xref="taxon:51511"
/clone="csef01lc15"
/dev_stage="egg"
/clone_lib="Yutaka Satou unpublished cDNA library (csefl)"

ORIGIN

Query Match 67.7%; Score 21; DB 5; Length 633;
Best Local Similarity 82.8%; Pred. No. 2.1e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTTGCGGCTTGGAA 31
|||||
Db 195 CATCGGCATCGTCATTTTCGGGTTGGTAC 223
|||||

RESULT 6
BB093704
LOCUS
DEFINITION
ACCSSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BB093704 RIKEN full-length enriched, 12 days embryo, embryonic body
between diaphragm region and neck Mus musculus cDNA clone
9430043J03 3' similar to AF191018 Homo sapiens E2IG3 (E2IG3) mRNA,
mRNA sequence.
BB093704
BB093704.2 GI:16261146
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 656)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Kono, H., Kouda, M., Koyama, S., Matsuura, T., Miyazaki, A., Nomura, K.,
Onno, H., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
On Jun 23, 2000 this sequence version replaced gi:8676953.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10
(11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.


```

Query Match      67.7%; Score 21; DB 4; Length 847;
Best Local Similarity 82.8%; Pred. No. 2.1e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CTCAACGGCATCGTCAGTTCGGGCTTGA 29
    |||||
Db 112 CTGAAGGCATCCTCAGTTCGGGCTTGA 140

RESULT 9
BU516442
LOCUS
DEFINITION AGNCOURT 10105175 NIH_MGC_134 Mus musculus cDNA clone
IMAGE:6513169 5', mRNA sequence.
ACCESSION BU516442
VERSION BU516442.1 GI:22823968
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 975)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14086 row: h column: 02
High quality sequence stop: 649.

FEATURES
source
1..975
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/c clones="IMAGE:6513169"
/tissue_type="undifferentiated limb"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_134"
/notes="vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match      67.7%; Score 21; DB 5; Length 975;
Best Local Similarity 82.8%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CTCAACGGCATCGTCAGTTCGGGCTTGA 29
    |||||
Db 665 CTGAAGGCATCCTCAGTTCGGGCTTGA 693

RESULT 10
AK034825
LOCUS
DEFINITION Mus musculus 12 days embryonic body between diaphragm region
and neck cDNA, RIKEN full-length enriched library, clone:9430043J03
product:similar to PUTATIVE NUCLEOTIDE BINDING PROTEIN,
ESTRADIOL-INDUCED (Homo sapiens), full insert sequence.
ACCESSION AK034825
VERSION AK034825.1 GI:26084222
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itch,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED 11042159
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagao,K., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED 11076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1653)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohashi,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1..1653
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

/db_xref="FANTOM DB:9430043J03"
 /db_xref="taxon:10090"
 /clone="9430043J03"
 /tissue_type="embryonic body between diaphragm region and neck"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"

misc_feature

1..1653
 /note="similar to PUTATIVE NUCLEOTIDE BINDING PROTEIN, ESTRADIOL-INDUCED [Homo sapiens] (SPTR|Q9BVP2, evidence: FASTY, 71.9%ID, 100%length, match=1614)"

ORIGIN

Query Match 67.7%; Score 21; DB 3; Length 1653;
 Best Local Similarity 82.8%; Pred. No. 2.3e+02;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTCAAGGCATCGTCAGTTCGGCTTGA 29

Db 1332 CTGAAGGCATCTCAGTTCGGCTTGA 1360

RESULT 11

CK744057/c

LOCUS CK744057 516 bp mRNA linear EST 24-FEB-2004
 DEFINITION ltu01-14ms3-d11 Ltu01 Liriodendron tulipifera cDNA clone
 ltu01-14ms3-d11 5', mRNA sequence.

ACCESSION CK744057

VERSION CK744057

KEYWORDS EST.

SOURCE Liriodendron tulipifera

ORGANISM Liriodendron tulipifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; magnoliids; Magnoliales;
 Magnoliaceae; Liriodendron.

REFERENCE 1 (bases 1 to 516)

Oppenheimer, D., Frohlich, M., Doyle, J., Tanksley, S., Webb, M.,
 Leebens-Mack, J., Landherr, L., Schlarbaum, S., Ilut, D. and Wall, K.
 Generation of ESTs from early flower buds of Liriodendron

tulipifera

Unpublished (2003)

Contact: Claude deFamphilis or James Leebens-Mack

Mueller Laboratory

Penn State University

208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn

State University, University Park, PA 16802, USA

Tel: 814 863 6413

Fax: 814 865 9131

Email: cwd3@psu.edu or jhl10@psu.edu

The sequence provided is trimmed of vector and low quality regions.
 Full sequence and original trace file are available from the Plant
 Genome Network website (<http://pgn.cornell.edu>)

Plate: ltu01-14ms3 row: d column: 11

Seq primer: M13F.

FEATURES

source

1..516 Location/Qualifiers

/organism="Liriodendron tulipifera"

/mol_type="mRNA"

/db_xref="taxon:3415"

/clone="ltu01-14ms3-d11"

/tissue_type="flower buds"

/dev_stage="1-35 mm buds"

/lab_host="SOLR"

/clone_lib="ltu01"

/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
 Site 2: XhoI; This is a directionally cloned,
 non-normalized library. This library has been generated by
 the Floral Genome Project (FGP). The Floral Genome Project
 is funded by NSF's Plant Genome Research Program
 (DBI-0115684). More information about the project can be
 obtained at <http://fgp.bio.psu.edu>"

ORIGIN

Query Match 66.5%; Score 20.6; DB 7; Length 516;
 Best Local Similarity 85.2%; Pred. No. 3e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CAACGGCATCGTCAGTTCGGCTTGA 29

Db 433 CGACGGCATCGTTAATGGCGCTTGA 407

RESULT 12

CL155476/c

LOCUS CL155476

DEFINITION

ltu01-104 Sorghum bicolor genomic clone 10782170, genomic survey
 (LibID: 104) Sorghum bicolor genomic clone 10782170, genomic survey
 sequence.

ACCESSION CL155476

VERSION CL155476

KEYWORDS GSS.

SOURCE Sorghum bicolor (sorghum)

ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 493)

AUTHORS

Budiman, M.A., Flick, E., Jones, J., Nurnberg, A., Citek, R.W.,
 Robbins, D., Rohlfing, T., Bradford, K., Fries, J., McMenamy, J.,
 Trani, L., Isak, A., Zimmerman, C., Lakey, N. and Bedell, J.A.

Genethresher methylation filtered genomic sequences from Sorghum

bicolor

Unpublished (2004)

Contact: Bedell JA

Orion Genomics LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 341 row: 1 column: 02

Seq primer: M13/pUC Forward

Class: shotgun

High quality sequence stop: 493.

FEATURES

source

1..493 Location/Qualifiers

/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/cultivar="ATx623"

/db_xref="taxon:4558"

/clone="10782170"

/clone_lib="Sorghum methylation-filtered library (LibID: 104)"

/note="Organ: leaf; Vector: pBCK(-); Site 1: HincII; DNA

prepared from purified nuclei was randomly sheared,

end-repaired, size fractionated to enrich for the 0.5 to 5

kb fraction, ligated into HincII-digested pBCK(-) vector

and electroporated into E. coli cells. This is a

methylation-filtered library."

ORIGIN

Query Match 65.8%; Score 20.4; DB 9; Length 493;

Best Local Similarity 80.0%; Pred. No. 3.7e+02;

Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CTCAACGGCATCGTCAGTTCGGCTTGA 30

Db 230 CTCAACACGCCGTCAGTTCGGCTTGA 201

RESULT 13

BM401667/c

LOCUS BM401667

DEFINITION

J12H07F Snake Bothrops insularis library II3 Bothrops insularis

cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

ACCESSION BM401667


```

VERSION      BM401667.1  GI:20376295
KEYWORDS     EST.
SOURCE       Bothrops insularis (island jararaca)
ORGANISM     Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Bothrops.
REFERENCE    1 (bases 1 to 510)
AUTHORS      Junqueira-de-Azevedo, I.L.M. and Ho, P.L.
TITLE        A survey of gene expression and diversity in the venom glands of
the pitviper snake Bothrops insularis through the generation of
expressed sequence tags (ESTs)
JOURNAL      Gene 299 (1-2), 279-291 (2002)
MEDLINE      22347338
PUBMED       12459276
COMMENT      Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL01A (see Reference)
Seq primer: M13F.

FEATURES     Location/Qualifiers
             1..510
             /organism="Bothrops insularis"
             /mol_type="mRNA"
             /db_xref="taxon:8723"
             /tissue_type="venom glands"
             /clone_lib="Snake Bothrops insularis library IL3"
             /note="Organ: venom glands; Vector: pGEM1zf+; Site 1: Eco
RI; Site 2: Not I; Sug of mRNA from Bothrops insularis
venom glands were primed with oligo-(dT) and reverse
transcribed to cDNA using Superscript Plasmid System for
cDNA Synthesis and Cloning (Life Technologies). The cDNAs
were selected by size (350-600 pb and up 600 pb) in
agarose gel electrophoresis, linked to Eco RI adapters and
directionally cloned in pGEM1zf+ vector (Promega). ESTs
were generated from random clones and grouped in unique
sequences. The putative identification of each EST or
cluster was obtained through Blast searches (e-value <
e-05)."

ORIGIN
Query Match      65.8%; Score 20.4; DB 4; Length 510;
Best Local Similarity 77.4%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CTCACGGCATCGTCAGTTGGGCTTGGAAAC 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 CTCAGCTCATCTTCAGATTAGGCTTGNAAAC 451

RESULT 14
BM344014
LOCUS      BM344014
DEFINITION rr45f08.y1 Globodera rostochiensis J2 pcDNAII Smant v1 Globodera
rostochiensis cDNA 5', similar to SW:IF6 CAEEL O62106 EUKARYOTIC
TRANSLATION INITIATION FACTOR 6 ; mRNA sequence.
ACCESSION   BM344014
VERSION     BM344014.1  GI:18080929
KEYWORDS    EST.
SOURCE      Globodera rostochiensis
ORGANISM    Globodera rostochiensis
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida;
Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
REFERENCE    1 (bases 1 to 595)
AUTHORS      McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
Tsagarieishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

BM401667.1  GI:20376295
TITLE        Bothrops insularis (island jararaca)
JOURNAL      Unpublished (1999)
COMMENT      Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was contributed by Dr. Geert Smant of the Laboratory of
Nematology at Wageningen University, Wageningen,
Netherlands(geert.smant@emna.dpw.wau.nl). DNA Sequencing by:
Washington University Genome Sequencing Center
Seq primer: -40Rp from Gibco
High quality sequence stop: 448.

FEATURES     Location/Qualifiers
             1..595
             /organism="Globodera rostochiensis"
             /mol_type="mRNA"
             /db_xref="taxon:31243"
             /dev_stage="J2"
             /lab_host="DH10B"
             /clone_lib="Globodera rostochiensis J2 pcDNAII Smant v1"
             /note="Vector: pcDNAII (Invitrogen); Site 1: BstXI;
Site 2: EcoRI; The library was donated for sequencing by
Geert Smant from Wageningen University, Laboratory of
Nematology, The Netherlands."

ORIGIN
Query Match      65.8%; Score 20.4; DB 4; Length 595;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 TCAACGGCATCGTCAGTTGGGCTTGGAAAC 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 TCGCCGGCAGCGTCATCGCGGCTCGGAAC 548

RESULT 15
CA077416
LOCUS      CA077416
DEFINITION SCQSAM1031A10.b AM1 Saccharum officinarum cDNA clone SCQSAM1031A10
3', mRNA sequence.
ACCESSION   CA077416
VERSION     CA077416.1  GI:34929688
KEYWORDS    EST.
SOURCE      Saccharum officinarum
ORGANISM    Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE    1 (bases 1 to 609)
AUTHORS      Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
TITLE        The libraries that made SUCEST
JOURNAL      Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT      Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 031 row: A column: 10
Seq primer: SP6 Promoter primer.

FEATURES     Location/Qualifiers
             1..609
             /organism="Saccharum officinarum"

```

ORIGIN

Query Match	65.8%	Score 20.4;	DB 6;	Length 816;
Best Local Similarity	80.0%;	Pred. No. 3.9e+02;		
Matches 24;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;

Qy 2 TCAACGGCATCGTCAGTTGGCGCTTGGAAAC 31
||| ||||| ||||| ||||| ||||| |||||

Db 94 TCCTCGTCAGTCAGTTGGANTGGTAC 123

RESULT 18
CF885756/c

LOCUS
DEFINITION
tric083xh07.b11 T.reesei mycelial culture, Version 6 October 2003
ACCESSION
Hypocrea jecorina cDNA clone tric083xh07, mRNA sequence.
VERSION
CF885756
KEYWORDS
CF885756.1 GI:38140438
SOURCE
EST.
ORGANISM
Hypocrea jecorina (anamorph: Trichoderma reesei)
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocrea; Hypocrea.
REFERENCE
1 (bases 1 to 934)
AUTHORS
Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,
Teunissen, J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K.,
Ward, M. and Dean, R.A.
TITLE
Characterization of the protein processing and secretion pathways
in a comprehensive set of expressed sequence tags from Trichoderma
reesei
JOURNAL
FEMS Microbiol. Lett. 230 (2), 275-282 (2004)
COMMENT
Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph_dean@ncsu.edu
Seq primer: LT-F1 primer.
FEATURES
Location/Qualifiers
1..934
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric083xh07"
/dev_stage="mycelia"
/clone_l1b="T.reesei mycelial culture, Version 6 October
2003"
/notes="Vector: pREP3Y; Site_1: Not 1/sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
ORIGIN
Query Match 65.8%; Score 20.4; DB 7; Length 934;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 TCAACGGCATCGTCAGTTCGGCTTGGAC 31
Db 636 TTACCGCTTCGTCAGTTCGGCTTGGAC 607
RESULT 19
CNS07E9S/c
LOCUS
CNS07E9S 1151 bp DNA linear GSS 08-JUL-2001
DEFINITION
T3 end of clone XBD0AA002D03 of library XBD0AA from strain CBS 94
of Candida tropicalis, genomic survey sequence.
ACCESSION
AL441494
VERSION
AL441494.1 GI:12224720
KEYWORDS
GSS
SOURCE
Candida tropicalis
Candida tropicalis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
1 (bases 1 to 1151)
AUTHORS
Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Wincker, P. and Weissbach, J.
TITLE
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular association studies
JOURNAL
FEMS Lett. 487 (1), 3-12 (2000)
MEDLINE
20584711

PUBMED	11152876	
REFERENCE	2 (bases 1 to 1151)	
AUTHORS	Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,P. and Dujon,B.	
TITLE	Genomic exploration of the hemiascomycetous yeasts: 16. Candida tropicalis	
JOURNAL	FEBS Lett. 487 (1), 91-94 (2000)	
MEDLINE	20584726	
PUBMED	11152891	
REFERENCE	3 (bases 1 to 1151)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.	
FEATURES	Location/Qualifiers	
source	1..1151	
	/organism="Candida tropicalis"	
	/mol_type="genomic DNA"	
	/strain="CBS 94"	
	/db_xref="taxon:5482"	
	/clone="XBDOAA002D03"	
	/clone_lib="XBDOAA"	
	/note="end : T3"	
misc_feature	complement(<716..>967)	
	/note="similar to Saccharomyces cerevisiae ORF YBR212w [NGR1 ; glucose-repressible RNA-binding protein]"	
misc_feature	/evidence="not_experimental"	
	complement(<725..>1057)	
	/note="similar to Saccharomyces cerevisiae ORF YHR086w [NAM8 ; meiotic recombination protein]"	
	1 putative frameshift(s)"	
	/evidence="not_experimental"	
ORIGIN		
Query Match	65.8%;	Score 20.4; DB 9; Length 1151;
Best Local Similarity	80.1%;	Pred. No. 4.1e-02;
Matches	24; Conservative	0; Mismatches 6; Indels 0; Gaps 0;
QY	1	CTCAACGGCATCGTCAGTTGCGGCTTGAA 30
Db	540	CTCAACGGCATCATCATTTGGACCTGGA 511
RESULT 20		
AZ214463		
LOCUS	AZ214463	271 bp DNA linear GSS 09-JUN-2000
DEFINITION	Sheared DNA-104B7.TP Sheared DNA Trypanosoma brucei genomic clone	
ACCESSION	AZ214463	
VERSION	AZ214463.1	GI:8432263
KEYWORDS	GSS.	
SOURCE	Trypanosoma brucei	
ORGANISM	Trypanosoma brucei	
	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;	
REFERENCE	1 (bases 1 to 271)	
AUTHORS	El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C., Gerrard,C., Leach,V., de Jong,P., Ullu,E., Melville,S., Donelson,J., Fraser,C. and Adams,M.	
TITLE	Determination of clone end sequences from Trypanosoma brucei GUTat 10.1 sheared DNA library	
JOURNAL	unpublished (1999)	


```

Db      247 CTACGGCATGTCGCTTGGCGCTTG 223
      ||||||| ||| |||||||||
      337 bp  DNA  linear  GSS 29-JAN-2003
BX214909  BX214909  Dario rerio genomic clone DKEY-256F17, genomic survey sequence.
DEFINITION  BX214909
ACCESSION  BX214909
VERSION    BX214909.1  GI:28046795
KEYWORDS   GSS.
SOURCE     Dario rerio (zebrafish)
ORGANISM   Dario rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 337)
            Humphray,S.J., Huckle,E. and Durham,J.L.
            Direct Submission
            Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humphray@sanger.ac.uk Unpublished
            This sequence was generated from the T7 end of BAC 256F17. 256F17
            is part of the Daniokey BAC Library created by R. Plaasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source           1..337
                     /organism="Dario rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-256F17"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"

ORIGIN
Query Match      64.5%; Score 20; DB 9; Length 337;
Best Local Similarity 82.1%; Pred. No. 5.2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  1  CTCAACGGCATGTCAGTTCGGCTTGG 28
      ||||||| ||| |||||||||
      345 bp  mRNA  linear  EST 27-MAY-1998
AA985360  am30h08.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone
DEFINITION  IMAGE:1629471 3', mRNA sequence.
ACCESSION  AA985360
VERSION    AA985360.1  GI:3163885
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 345)
            Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
            Unpublished (1997)
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham.

TITLE      WashU-NCI human EST Project
JOURNAL    Unpublished (1997)
COMMENT    Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham.

FEATURES             Location/Qualifiers
     source           1..381
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1629471"
                     /sex="male"
                     /tissue_type="schizophrenic brain S-11 frontal lobe"
                     /dev_stage="34 years old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene schizo brain S11"

FEATURES             Location/Qualifiers
     source           1..381
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1629471"
                     /sex="male"
                     /tissue_type="schizophrenic brain S-11 frontal lobe"
                     /dev_stage="34 years old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene schizo brain S11"

```

/note=Vector: Bluescript SK-; Site 1: EcoRI; Library constructed from S-11 frontal lobe, male, 34 years old, 50% caucasian, 50% Aleutian. Schizophrenic suicide. Random primed into EcoRI site of ZAP II Vector. Mass excised. Custom library. Avg insert length 1.4kb. Material obtained by Johnston N., Torrey, E.F., Yolken R., and the Stanley Neuropathology Consortium - Analysis of RNAs from the Brains of Individuals with Psychiatric Diseases (Unpublished) Stanley Neurovirology Laboratory, Johns Hopkins School of Medicine, Baltimore MD."

ORIGIN

Query Match 64.5%; Score 20; DB 1; Length 381;
 Best Local Similarity 82.1%; Pred. No. 5.3e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 2 TCAACGGCAGTCAGTTGCGGCTTGA 29
 ||||| ||||| ||||| ||||| |||||
 Db 262 TCAAGGGAACGTCAGTTGCTGATTGA 289

Search completed: March 11, 2005, 13:01:10
 Job time : 413.213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 45.3508 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-13
Perfect score: 31
Sequence: 1 agcactcaacggcatcgctcagttgcggcttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Geneseqn16Dec04:*
2: Geneseqn1980s:*
3: Geneseqn1990s:*
4: Geneseqn2000s:*
5: Geneseqn2001as:*
6: Geneseqn2001bs:*
7: Geneseqn2002as:*
8: Geneseqn2002bs:*
9: Geneseqn2003as:*
10: Geneseqn2003bs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	31	AAZ36113	AAZ36113 Primer de
2	31	100.0	1489	AAZ36101	AAZ36101 Nucleic a
3	27	87.1	31	AAZ36112	AAZ36112 Primer de
C 4	20.6	66.5	1349	ABL04091	ABL04091 Drosophil
C 5	20.2	65.2	664	ADB68745	ADB68745 Pseudomon
C 6	20.2	65.2	1156	ACA48626	ACA48626 Prokaryot
C 7	20.2	65.2	1662	ACA50970	ACA50970 Prokaryot
C 8	20	64.5	1632	ADS80246	ADS80246 Bacterial
C 9	19.4	62.6	2043	ABQ90098	ABQ90098 M. capsul
10	19.4	62.6	89378	ADN47591_20	Continuation (21 o
11	19.4	62.6	89378	ADN47209_20	Continuation (21 o
12	19.4	62.6	89378	ADN47960_20	Continuation (21 o
C 13	19.4	62.6	110000	ADN46845_00	Adn46845 Thermococ
C 14	19.4	62.6	110000	ADN46123_00	Adn46123 Thermococ
C 15	19.4	62.6	110000	ADN46484_00	Adn46484 Thermococ
C 16	19.2	61.9	1278	ADN03059	Adn03059 Corn orth
C 17	19.2	61.9	1278	ADN42860	Adn42860 Plant tra
C 18	19.2	61.9	1362	ADT46111	Adt46111 Bacterial
C 19	19.2	61.9	1686	ACA51630	ACA51630 Prokaryot
C 20	19	61.3	1770	ADL03792	ADL03792 DNA encod

C 21	19	61.3	23210	4	AAF28530	Genomic f
C 22	19	61.3	41599	4	AAI66165	Bacillus
C 23	18.8	60.6	1275	11	ACH94680	Klebsiell
C 24	18.8	60.6	2571	4	ABL23899	Drosophil
C 25	18.8	60.6	4571	4	ABL23898	Drosophil
C 26	18.8	60.6	31477	11	ACN44096	Mouse gen
C 27	18.6	60.0	917	6	ABK72996	Bacillus
C 28	18.6	60.0	1056	13	ADS46729	Bacterial
C 29	18.6	60.0	1350	5	AAS80533	DNA encod
C 30	18.6	60.0	2556	5	AAS80531	DNA encod
C 31	18.6	60.0	2560	5	AAS94344	DNA encod
C 32	18.6	60.0	3150	4	AAS52338	E. coli D
C 33	18.6	60.0	3150	8	ACA19089	Prokaryot
C 34	18.6	60.0	4879	2	AAZ11064	E. coli a
C 35	18.6	60.0	9921	11	ADO59400	Antheraea
C 36	18.4	59.4	552	11	ABD17546	Pseudomon
C 37	18.4	59.4	576	11	ACH95920	Klebsiell
C 38	18.4	59.4	636	6	AAL49962	Human B11
C 39	18.4	59.4	824	6	AAL49970	Human B11
C 40	18.4	59.4	1018	6	ABV77583	Glycerald
C 41	18.4	59.4	1077	4	AAS54336	Pseudomon
C 42	18.4	59.4	1077	8	ACA42699	Prokaryot
C 43	18.4	59.4	1107	11	ABD17845	Pseudomon
C 44	18.4	59.4	1143	9	ADB80402	Human MDD
C 45	18.4	59.4	1410	11	ABD17735	Pseudomon

ALIGNMENTS

RESULT 1

AAZ36113 ID AAZ36113 standard; DNA; 31 BP.

XX AC AAZ36113;

DT 11-FEB-2000 (first entry)

DE Primer derived from a nucleic acid sequence specific to EHEC.

OS Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
PCR primer; probe; ss.

XX OS Synthetic.

OS Escherichia coli.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
Escherichia coli, particularly serotype O157:H7, used for detecting these
bacteria in food.

XX PS Claim 5; Page 27; 48pp; French.

XX CC AAZ36103-27 represent fragments derived from nucleic acid sequences
specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
derived from two sequences. The first (AAZ36101) is 99.9% homologous to
the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies

XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCACTCAACGGCATCTCAGTTGCGGCTTG 31
| | | | | | | | | | | | | | | | | | | | |
Db 1 AGCACTCAACGGCATCTCAGTTGCGGCTTG 31

RESULT 2
AAZ36101
ID AAZ36101 standard; DNA; 1489 BP.

XX AC AAZ36101;

XX DT 11-FEB-2000 (first entry)

XX DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX IS91; ds.

XX OS Escherichia coli.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
XX Escherichia coli, particularly serotype O157:H7, used for detecting these
XX bacteria in food.

XX PS Claim 1; Fig 1; 48pp; French.

XX CC The present sequence is specific to enterohemorrhagic Escherichia coli
XX (EHEC). The sequence is 99.9% homologous to the katP gene of E. coli
XX O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
XX homologous with IS91 of E. coli (nucleotides 1-406 of the present
XX sequence). The present sequence is of plasmid origin. Fragments of the
XX present sequence are used, as probes and primers, for detection of E.
XX coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
XX animal samples, foods or the environment. The fragments are also useful
XX for epidemiological studies

XX SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 1489;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCACTCAACGGCATCTCAGTTGCGGCTTG 31
| | | | | | | | | | | | | | | | | | | | |
Db 387 AGCACTCAACGGCATCTCAGTTGCGGCTTG 417

RESULT 3

AAZ36112
ID AAZ36112 standard; DNA; 31 BP.

XX AC AAZ36112;

XX DT 11-FEB-2000 (first entry)

XX DE Primer derived from a nucleic acid sequence specific to EHEC.

XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
XX PCR primer; probe; ss.

XX OS Synthetic.

XX OS Escherichia coli.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
XX Escherichia coli, particularly serotype O157:H7, used for detecting these
XX bacteria in food.

XX PS Claim 5; Page 27; 48pp; French.

XX CC AAZ36103-27 represent fragments derived from nucleic acid sequences
XX specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
XX derived from two sequences. The first (AAZ36101) is 99.9% homologous to
XX the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
XX 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
XX The second sequence (AAZ36102) is associated with the presence of
XX virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
XX 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
XX for virulence proteins of Shigella flexneri. Both sequences are of
XX plasmid origin. The fragments are used as PCR primers and probes for the
XX detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
XX in human or animal samples, foods or the environment. The fragments are
XX also useful for epidemiological studies

XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 87.1%; Score 27; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 CTCACGGCATCTCAGTTGCGGCTTG 31
| | | | | | | | | | | | | | | | | | | | |
Db 1 CTCACGGCATCTCAGTTGCGGCTTG 27

RESULT 4

ABL04091/c
ID ABL04091 standard; cDNA; 1349 BP.

XX AC ABL04091;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 6755.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; Gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li FWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB5998.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 6755; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1349 BP; 337 A; 366 C; 351 G; 295 T; 0 U; 0 Other;
Query Match 66.5%; Score 20.6; DB 4; Length 1349;
Best Local Similarity 85.2%; Pred. No. 70;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1 AGCACTCAACGGCATCGTCAGTTCGG 27
Db 872 AGCCCTTAATGGCATCATCAGTTCGG 846
RESULT 5
ADB68745/c
ID ADB68745 standard; DNA; 664 BP.
XX AC ADB68745;
XX 04-DEC-2003 (first entry)
XX Pseudomonas chlororaphis phzI DNA.
XX quorum sensing; lux homologue; luxI; ds.
XX Pseudomonas chlororaphis.
XX WO2003057902-A2.
XX 17-JUL-2003.
XX 08-JAN-2003; 2003WO-US000479.
XX 08-JAN-2002; 2002US-0346531P.
XX 07-JAN-2003; 2003US-00338110.
XX

PA (FEAU) FRAUNHOFER USA INC.
PI Fuhrmann JJ, Romesser JA;
XX WPI; 2003-618102/58.
XX Detecting quorum sensing potential of a Gram-negative bacterium in a
PT sample comprises performing a polymerase chain reaction using nucleic
PT acids extracted from a sample containing a microorganism.
XX Disclosure; Fig 10; 86pp; English.
XX The invention relates to a novel method for detecting the quorum sensing
CC potential of a microorganism in a sample which comprises performing PCR
CC using nucleic acids extracted from a sample containing at least one type
CC of microorganism. The method may be useful for detecting the quorum
CC sensing potential of a microorganism in a sample by amplifying a fragment
CC of a lux gene or homologue. The current sequence is that of the luxI
CC homologue DNA of the invention.
XX Sequence 664 BP; 162 A; 219 C; 167 G; 116 T; 0 U; 0 Other;
Query Match 65.2%; Score 20.2; DB 10; Length 664;
Best Local Similarity 88.0%; Pred. No. 94;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 7 CAACGGCATCGTCAGTTCGGCTTG 31
Db 421 CAACGGCATCGCCAGTTCGGCTTG 397
RESULT 6
ACA48626/c
ID ACA48626 standard; DNA; 1156 BP.
XX AC ACA48626;
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #30283.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX Salmonella paratyphi.
OS WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU44756.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 36496; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of

PA (FEAU) FRAUNHOFER USA INC.
PI Fuhrmann JJ, Romesser JA;
XX WPI; 2003-618102/58.
XX Detecting quorum sensing potential of a Gram-negative bacterium in a
PT sample comprises performing a polymerase chain reaction using nucleic
PT acids extracted from a sample containing a microorganism.
XX Disclosure; Fig 10; 86pp; English.
XX The invention relates to a novel method for detecting the quorum sensing
CC potential of a microorganism in a sample which comprises performing PCR
CC using nucleic acids extracted from a sample containing at least one type
CC of microorganism. The method may be useful for detecting the quorum
CC sensing potential of a microorganism in a sample by amplifying a fragment
CC of a lux gene or homologue. The current sequence is that of the luxI
CC homologue DNA of the invention.
XX Sequence 664 BP; 162 A; 219 C; 167 G; 116 T; 0 U; 0 Other;
Query Match 65.2%; Score 20.2; DB 10; Length 664;
Best Local Similarity 88.0%; Pred. No. 94;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 7 CAACGGCATCGTCAGTTCGGCTTG 31
Db 421 CAACGGCATCGCCAGTTCGGCTTG 397
RESULT 6
ACA48626/c
ID ACA48626 standard; DNA; 1156 BP.
XX AC ACA48626;
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #30283.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX Salmonella paratyphi.
OS WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU44756.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 36496; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1156 BP; 196 A; 320 C; 341 G; 297 T; 0 U; 2 Other;

Query Match 65.2%; Score 20.2; DB 8; Length 1156;
 Best Local Similarity 88.0%; Pred. No. 1e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ACTCACGGCATCGTCAGTTGCGGC 28
 ||||| ||||| ||||| ||||| |||||
 Db 294 ACTCCACGGCAGCATCAGTTGCGGC 270

RESULT 7
 ACA50970/c
 ID ACA50970 standard; DNA; 1662 BP.

XX AC ACA50970;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #32627.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX drug design; gene.

XX *Salmonella typhimurium*.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 08-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone G, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI: 2003-029926/02.

DR P-PSDB; ABU47100.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 38840; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1662 BP; 274 A; 457 C; 492 G; 439 T; 0 U; 0 Other;

Query Match 65.2%; Score 20.2; DB 8; Length 1662;
 Best Local Similarity 88.0%; Pred. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ACTCACGGCATCGTCAGTTGCGGC 28

Db 762 ACTCCACGGCAGCATCAGTTGCGGC 738
 ||||| ||||| ||||| ||||| |||||

RESULT 8

ADS60246/c

ID ADS60246 standard; cDNA; 1632 BP.

XX AC ADS60246;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #12233.

XX Recombinant DNA construct; transformed plant; improved plant property;
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 XX pathogen tolerance; pest tolerance; plant disease resistance;
 XX cell cycle pathway modification; plant growth regulator;
 XX homologous recombination; seed oil yield; protein yield; carbohydrate;
 XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 XX bacterial polynucleotide; gene; ss.

XX Bacteria.

XX

PN US2003233675-A1.
XX 18-DEC-2003.
PD 20-FEB-2003; 2003US-00369493.
PF 21-FEB-2002; 2002US-0360039P.
PP (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 35920; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polynucleotide used in
XX the scope of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX SQ Sequence 1632 BP; 299 A; 506 C; 545 G; 282 T; 0 U; 0 Other;
Query Match 64.5%; Score 20; DB 13; Length 1632;
Best Local Similarity 82.1%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 ACTCAACGGATCGTCAGTTGGCGTTG 31
DB 804 ACCGACAGCATCATCAGTTGGCGGTG 777
RESULT 9
ABQ90098/C
ID ABQ90098 standard; DNA; 2043 BP.
XX AC ABQ90098;
XX DT 01-OCT-2002 (first entry)
XX XX M. capsulatus gene #83 for DNA array.
XX DE Micro array; gene; ds; differential expression; gene expression.
XX KW Methylcoccus capsulatus.
XX OS

XX WO200255655-A2.
XX 18-JUL-2002.
XX 14-JAN-2002; 2002WO-NO000019.
XX 12-JAN-2001; 2001NO-00000235.
XX 12-JAN-2001; 2001NO-00000239.
XX (UNIF-) UNIFOB STIFTTELSEN UNIV BERGEN.
XX (TIGR-) TIGR.
XX Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
XX Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
XX Salzberg SL;
XX WPI; 2002-557818/59.
XX Novel DNA array useful for determining differential expression of
XX Methylcoccus capsulatus genes, comprises polynucleotides or
XX oligonucleotides representative for a selective number of Methylcoccus
XX capsulatus genes.
XX Claim 19; Page 90; 678pp; English.
XX The invention relates to a novel DNA array giving a representation of a
XX number of Methylcoccus capsulatus genes. The method of the invention is
XX useful for determination of the differential expression of the genes of
XX M. capsulatus, and for studying gene expression on a genomic scale and in
XX gene expression assays of M. capsulatus genes. The sequences shown in
XX ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the
XX invention
XX SQ Sequence 2043 BP; 395 A; 700 C; 648 G; 300 T; 0 U; 0 Other;
Query Match 62.6%; Score 19.4; DB 6; Length 2043;
Best Local Similarity 79.3%; Pred. No. 2.5e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 GCATCTCAACGCGATCGTCAGTTGGCGTT 30
DB 1638 GGACTCGACGCGATCGCGGTTTCGGGTT 1610
RESULT 10
ADN47591_20
Continuation (21 of 21) of ADN47591 from base 2000001 (Thermococcus kodakaraensis KOD1 9
WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591
WP Fragment Name Begin End
WP ADN47591_00 1 110000
WP ADN47591_01 100001 210000
WP ADN47591_02 200001 310000
WP ADN47591_03 300001 410000
WP ADN47591_04 400001 510000
WP ADN47591_05 500001 610000
WP ADN47591_06 600001 710000
WP ADN47591_07 700001 810000
WP ADN47591_08 800001 910000
WP ADN47591_09 900001 1010000
WP ADN47591_10 1000001 1110000
WP ADN47591_11 1100001 1210000
WP ADN47591_12 1200001 1310000
WP ADN47591_13 1300001 1410000
WP ADN47591_14 1400001 1510000
WP ADN47591_15 1500001 1610000
WP ADN47591_16 1600001 1710000
WP ADN47591_17 1700001 1810000
WP ADN47591_18 1800001 1910000
WP ADN47591_19 1900001 2010000
WP ADN47591_20 2000001 2089378
Query Match 62.6%; Score 19.4; DB 12; Length 89378;

Best Local Similarity 79.3%; Pred. No. 4.3e+02;		Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	2	GCACCTCAACGGCATCGTCAGTTCGGGCTT 30	
Db	32709	GCACCTCAACTCCCTCCTCAGTTGATGCTT 32737	
RESULT 11			
ADN47209_20			
Continuation (21 of 21) of ADN47209 from base 2000001 (Thermococcus kodakaraensis KOD1 g			
WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209			
WP	Fragment Name	Begin	End
WP	ADN47209_00	1	110000
WP	ADN47209_01	100001	210000
WP	ADN47209_02	200001	310000
WP	ADN47209_03	300001	410000
WP	ADN47209_04	400001	510000
WP	ADN47209_05	500001	610000
WP	ADN47209_06	600001	710000
WP	ADN47209_07	700001	810000
WP	ADN47209_08	800001	910000
WP	ADN47209_09	900001	1010000
WP	ADN47209_10	1000001	1110000
WP	ADN47209_11	1100001	1210000
WP	ADN47209_12	1200001	1310000
WP	ADN47209_13	1300001	1410000
WP	ADN47209_14	1400001	1510000
WP	ADN47209_15	1500001	1610000
WP	ADN47209_16	1600001	1710000
WP	ADN47209_17	1700001	1810000
WP	ADN47209_18	1800001	1910000
WP	ADN47209_19	1900001	2010000
WP	ADN47209_20	2000001	2089378
Query Match 62.6%; Score 19.4; DB 12; Length 89378;			
Best Local Similarity 79.3%; Pred. No. 4.3e+02;		Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	2	GCACCTCAACGGCATCGTCAGTTCGGGCTT 30	
Db	32709	GCACCTCAACTCCCTCCTCAGTTGATGCTT 32737	
RESULT 12			
ADN47960_20			
Continuation (21 of 21) of ADN47960 from base 2000001 (Thermococcus kodakaraensis KOD1 g			
WP Sequence split into 21 fragments LOCUS ADN47960 Accession Adn47960			
WP	Fragment Name	Begin	End
WP	ADN47960_00	1	110000
WP	ADN47960_01	100001	210000
WP	ADN47960_02	200001	310000
WP	ADN47960_03	300001	410000
WP	ADN47960_04	400001	510000
WP	ADN47960_05	500001	610000
WP	ADN47960_06	600001	710000
WP	ADN47960_07	700001	810000
WP	ADN47960_08	800001	910000
WP	ADN47960_09	900001	1010000
WP	ADN47960_10	1000001	1110000
WP	ADN47960_11	1100001	1210000
WP	ADN47960_12	1200001	1310000
WP	ADN47960_13	1300001	1410000
WP	ADN47960_14	1400001	1510000
WP	ADN47960_15	1500001	1610000
WP	ADN47960_16	1600001	1710000
WP	ADN47960_17	1700001	1810000
WP	ADN47960_18	1800001	1910000
WP	ADN47960_19	1900001	2010000
WP	ADN47960_20	2000001	2089378
Query Match 62.6%; Score 19.4; DB 12; Length 89378;			
Best Local Similarity 79.3%; Pred. No. 4.3e+02;		Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	

Best Local Similarity 79.3%; Pred. No. 4.3e+02;		Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	2	GCACCTCAACGGCATCGTCAGTTCGGGCTT 30	
Db	32709	GCACCTCAACTCCCTCCTCAGTTGATGCTT 32737	
RESULT 13			
ADN46845_00/c			
WP Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845			
WP	Fragment Name	Begin	End
WP	ADN46845_00	1	110000
WP	ADN46845_01	100001	210000
WP	ADN46845_02	200001	310000
WP	ADN46845_03	300001	410000
WP	ADN46845_04	400001	510000
WP	ADN46845_05	500001	610000
WP	ADN46845_06	600001	710000
WP	ADN46845_07	700001	810000
WP	ADN46845_08	800001	910000
WP	ADN46845_09	900001	1010000
WP	ADN46845_10	1000001	1110000
WP	ADN46845_11	1100001	1210000
WP	ADN46845_12	1200001	1310000
WP	ADN46845_13	1300001	1410000
WP	ADN46845_14	1400001	1510000
WP	ADN46845_15	1500001	1610000
WP	ADN46845_16	1600001	1710000
WP	ADN46845_17	1700001	1810000
WP	ADN46845_18	1800001	1910000
WP	ADN46845_19	1900001	2010000
WP	ADN46845_20	2000001	2089378
ID	ADN46845 standard; DNA; 2089378 BP.		
XX	AC		
XX	ADN46845;		
XX	01-JUL-2004 (first entry)		
XX	Thermococcus kodakaraensis KOD1 genome DNA sequence SeqID723.		
DE	gene disruption; gene targeting; marker gene; transformation;		
XX	homologous recombination; hyperthermostable archaeobacterium; KOD1;		
KW	gene structure; gene function; enzyme activity; medicine;		
KW	forensic science; food; drug inspection; molecular biology; immunology;		
KW	ds.		
XX	Thermococcus kodakaraensis.		
OS	WO2004022736-A1.		
XX	18-MAR-2004.		
XX	29-AUG-2003; 2003WO-IB003597.		
XX	30-AUG-2002; 2002JP-00319011.		
XX	(NISC-) JAPAN SCI & TECHNOLOGY CORP.		
XX	Imanaka T, Atomi H;		
XX	WPI; 2004-257583/24.		
XX	Method for disrupting targeted gene in genome of organism particularly		
PT	thermostable bacterium and with genome chips for analysis, applicable in		
PT	studying gene structure and functions.		
XX	Example 1; SEQ ID NO 723; 598pp; Japanese.		
XX	This invention relates to a novel method for targeting disruption of an		
CC	arbitrary gene in a genome of an organism which comprises providing the		
CC	whole sequential data of the genome of such organism, selecting at least		
CC	1 arbitrary region in the sequence, providing a vector that contains a		
CC	sequence homologous with the selected region and a marker gene,		

CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakaraensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of the genomic DNA
CC sequence of Thermococcus kodakaraensis (KOD1) which was derived during
CC use of the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2089378 BP; 503594A; 544081C; 542308G; 499376T; 0U; 190ther;
Query Match 62.6%; Score 19.4; DB 12; Length 110000;
Best Local Similarity 79.3%; Pred. No. 4.4e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 GCACCTCAACGGCATCGTCAGTTGGCGCTT 30
Db 56669 GCACCTCAACCTCCCTCCTCAGTTGATGCTT 56641

RESULT 14
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123
WP Fragment Name Begin End
WP ADN46123_00 1 110000
WP ADN46123_01 100001 210000
WP ADN46123_02 200001 310000
WP ADN46123_03 300001 410000
WP ADN46123_04 400001 510000
WP ADN46123_05 500001 610000
WP ADN46123_06 600001 710000
WP ADN46123_07 700001 810000
WP ADN46123_08 800001 910000
WP ADN46123_09 900001 1010000
WP ADN46123_10 1000001 1110000
WP ADN46123_11 1100001 1210000
WP ADN46123_12 1200001 1310000
WP ADN46123_13 1300001 1410000
WP ADN46123_14 1400001 1510000
WP ADN46123_15 1500001 1610000
WP ADN46123_16 1600001 1710000
WP ADN46123_17 1700001 1810000
WP ADN46123_18 1800001 1910000
WP ADN46123_19 1900001 2010000
WP ADN46123_20 2000001 2089378
ID ADN46123 standard; DNA; 2089378 BP.

XX AC ADN46123;
XX 01-JUL-2004 (first entry)
DE Thermococcus kodakaraensis KOD1 genome DNA sequence SeqID1.
XX gene disruption; gene targeting; marker gene; transformation;
XX homologous recombination; hyperthermostable archaeobacterium; KOD1;
XX gene structure; gene function; enzyme activity; medicine;
XX forensic science; food; drug inspection; molecular biology; immunology;
XX ds; gene.
XX Thermococcus kodakaraensis.
XX WO2004022736-A1.
XX 18-MAR-2004.
XX 29-AUG-2003; 2003WO-IB003597.

PR 30-AUG-2002; 2002JP-00319011.
FA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Imanaka T, Atom1 H;
XX WPI; 2004-257583/24.
XX Method for disrupting targeted gene in genome of organism particularly
XX thermostable bacterium and with genome chips for analysis, applicable in
XX studying gene structure and functions.
PS Claim 8; SEQ ID NO 1; 598pp; Japanese.
XX This invention relates to a novel method for targeting disruption of an
XX arbitrary gene in a genome of an organism which comprises providing the
XX whole sequential data of the genome of such organism, selecting at least
XX 1 arbitrary region in the sequence, providing a vector that contains a
XX sequence homologous with the selected region and a marker gene
XX transformation, and homologous recombination. The genome is preferably
XX the genome of a hyperthermostable archaeobacterium, particularly
XX Thermococcus kodakaraensis KOD1. The method is for targeting the
XX disruption of a gene in the genome of an organism, which is applicable in
XX studying gene structure and functions as well as enzyme activities of
XX encoded proteins and useful in medicine, forensic science, food or drug
XX inspection, molecular biology and immunology. With this method, the
XX disruption of a gene at an arbitrary position in a genome can be achieved
XX efficiently and reliably. The present sequence is that of the genomic DNA
XX sequence of Thermococcus kodakaraensis (KOD1) which was derived during
XX use of the method of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2089378 BP; 503594A; 544081C; 542308G; 499376T; 0U; 190ther;
Query Match 62.6%; Score 19.4; DB 12; Length 110000;
Best Local Similarity 79.3%; Pred. No. 4.4e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 GCACCTCAACGGCATCGTCAGTTGGCGCTT 30
Db 56669 GCACCTCAACCTCCCTCCTCAGTTGATGCTT 56641

RESULT 15
WP Sequence split into 21 fragments LOCUS ADN46464 Accession Adn46464
WP Fragment Name Begin End
WP ADN46464_00 1 110000
WP ADN46464_01 100001 210000
WP ADN46464_02 200001 310000
WP ADN46464_03 300001 410000
WP ADN46464_04 400001 510000
WP ADN46464_05 500001 610000
WP ADN46464_06 600001 710000
WP ADN46464_07 700001 810000
WP ADN46464_08 800001 910000
WP ADN46464_09 900001 1010000
WP ADN46464_10 1000001 1110000
WP ADN46464_11 1100001 1210000
WP ADN46464_12 1200001 1310000
WP ADN46464_13 1300001 1410000
WP ADN46464_14 1400001 1510000
WP ADN46464_15 1500001 1610000
WP ADN46464_16 1600001 1710000
WP ADN46464_17 1700001 1810000
WP ADN46464_18 1800001 1910000
WP ADN46464_19 1900001 2010000
WP ADN46464_20 2000001 2089378
ID ADN46464 standard; DNA; 2089378 BP.
XX AC ADN46464;

```
XX 01-JUL-2004 (first entry)
XX Thermococcus kodakaraensis KOD1 DNA sequence SeqID342.
XX
XX gene disruption; gene targeting; marker gene; transformation;
XX homologous recombination; hyperthermostable archaeobacterium; KOD1;
XX gene structure; gene function; enzyme activity; medicine;
XX forensic science; food; drug inspection; molecular biology; immunology;
XX ds; gene.
XX
XX Thermococcus kodakaraensis.
XX
XX WO2004022736-A1.
XX
XX 18-MAR-2004.
XX
XX 29-AUG-2003; 2003WO-IB003597.
XX
XX 30-AUG-2002; 2002JP-00319011.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Imanaka T, Atomi H;
XX
XX WPI; 2004-257583/24.
XX
XX Method for disrupting targeted gene in genome of organism particularly
XX thermostable bacterium and with genome chips for analysis, applicable in
XX studying gene structure and functions.
XX
XX Example 1; SEQ ID NO 342; 598pp; Japanese.
XX
XX This invention relates to a novel method for targeting disruption of an
XX arbitrary gene in a genome of an organism which comprises providing the
XX whole sequential data of the genome of such organism, selecting at least
XX 1 arbitrary region in the sequence, providing a vector that contains a
XX sequence homologous with the selected region and a marker gene,
XX transformation, and homologous recombination. The genome is preferably
XX the genome of a hyperthermostable archaeobacterium, particularly
XX Thermococcus kodakaraensis KOD1. The method is for targeting the
XX disruption of a gene in the genome of an organism, which is applicable in
XX studying gene structure and functions as well as enzyme activities of
XX encoded proteins and useful in medicine, forensic science, food or drug
XX inspection, molecular biology and immunology. With this method, the
XX disruption of a gene at an arbitrary position in a genome can be achieved
XX efficiently and reliably. The present sequence is that of the genomic DNA
XX sequence of Thermococcus kodakaraensis (KOD1) which was derived during
XX use of the method of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2089378 BP; 503594A; 544081C; 542308G; 499376T; 0U; 190Other;
XX
XX Query Match 62.68; Score 19.4; DB 12; Length 110000;
XX Best Local Similarity 79.38; Pred. No. 4.4e+02;
XX Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 2 GCACCTAACCGCATCGTCAGTTGGCGCTT 30
XX
XX Db 56669 GCACCTAACCTCCCTCCCTCAGTTGATGCTT 56641
XX
XX
XX RESULT 16
XX ADO03059/c
XX ID ADO03059 standard; cDNA; 1278 BP.
XX
XX XX
XX AC ADO03059;
XX
XX 01-JUL-2004 (first entry)
XX
XX DE Corn orthologue of Thalecress transcription factor, cDNA #134.
```

```
XX
XX Corn; transcription factor; ss; gene; plant; transgenic; abiotic stress;
XX cold tolerance; heat tolerance; drought; osmotic stress;
XX phosphate limitation; potassium limitation; nitrogen limitation;
XX hormone sensitivity; disease resistance; sugar sensing; seed germination;
XX flowering; inflorescence architectural change;
XX meristem cell differentiation; phyllotaxy; apical dominance;
XX trichome development; seed development; premature senescence;
XX delayed senescence; lethality; necrosis; plant size; leaf morphology;
XX seed morphology; secondary metabolism; light response; shade avoidance.
XX
XX Zea mays.
XX
XX OS
XX US2004045049-A1.
XX
XX PN
XX 04-MAR-2004.
XX
XX PD
XX
XX PF
XX 10-APR-2003; 2003US-00412699.
XX
XX PR
XX 13-SEP-1999; 99US-00394519.
XX 21-JAN-2000; 2000US-00489376.
XX 17-FEB-2000; 2000US-00506720.
XX 22-MAR-2000; 2000US-00532591.
XX 22-MAR-2000; 2000US-00533029.
XX 22-MAR-2000; 2000US-00533030.
XX 22-MAR-2000; 2000US-00533392.
XX 22-MAR-2000; 2000US-00533648.
XX 06-APR-2000; 2000WO-US009448.
XX 16-NOV-2000; 2000US-00713994.
XX 27-MAR-2001; 2001US-00819142.
XX 17-APR-2001; 2001US-00837444.
XX 30-JAN-2002; 2002US-00958131.
XX 14-JUN-2002; 2002US-00171468.
XX 09-AUG-2002; 2002US-00225066.
XX 09-AUG-2002; 2002US-00225067.
XX 17-DEC-2002; 2002US-00225068.
XX 25-FEB-2003; 2002US-0434166P.
XX 25-FEB-2003; 2003US-00374780.
XX
XX (ZHAN/) ZHANG J.
XX (FROM/) FROMM M E.
XX (HEAR/) HEARD J E.
XX (RIEC/) RIECHMANN J L.
XX (ADAM/) ADAM L J.
XX (BROU/) BROUN P E.
XX (PINE/) PINEDA O.
XX (REUB/) REUBER T L.
XX (KEDD/) KEDDIE J S.
XX (YUGG/) YU G.
XX (JIAN/) JIANG C.
XX (SAMA/) SAMAHA R S.
XX (PILG/) PILGRIM M L.
XX (CREE/) CREELMAN R A.
XX (DUBE/) DUBELL A N.
XX (RATC/) RATCLIFFE O.
XX (KUMI/) KUMIMOTO R.
XX (SHER/) SHERMAN B K.
XX
XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
XX Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
XX Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
XX Sherman BK;
XX
XX WPI; 2004-225755/21.
XX
XX New transgenic plant, useful in developing phenotypes with altered or
XX improved characteristics or traits.
XX
XX Claim 1; SEQ ID NO 1473; 213pp; English.
XX
XX The invention relates to a transgenic plant comprises a recombinant
XX polynucleotide having a polynucleotide sequence or its complementary
XX sequence comprising a sequence encoding a polypeptide, that initiates
```

transcription (i.e. a transcription factor) from Arabidopsis, Soybean, Rice, Rape or Corn, comprising any of the sequences appearing as AD001588-AD003527 or AD003530-AD003559. Also included are using a transgenic plant to grow a progeny plant, an expression cassette (comprising a constitutive, inducible or tissue-specific promoter and a recombinant polynucleotide described above), a host cell comprising the expression cassette, producing a modified plant having a modified trait, identifying a factor that is modulated by or interacts with a polypeptide encoded by the polynucleotide sequence and identifying at least one downstream polynucleotide sequence that is subject to a regulatory effect of any of the polypeptides encoded by the polynucleotide described above. The transgenic plant is useful for producing a plant that has an altered trait e.g. an enhanced tolerance to abiotic stress (increased tolerance to chilling, germination in cold conditions, freezing tolerance, tolerance to heat, tolerance to drought, tolerance to osmotic stress, tolerance to salt, tolerance to phosphate limitation, tolerance to potassium limitation, decreased sensitivity to nitrogen limitation), altered hormone sensitivity, reduced sensitivity to abscisic acid, an altered response to ethylene, disease resistance, altered susceptibility to Botrytis, altered susceptibility to Fusarium, altered susceptibility to Erysiphe, altered susceptibility to Pseudomonas syringae, altered susceptibility to Sclerotinia, altered sugar sensing, improved seed germination and seedling vigor, early flowering, late flowering, extended period of flowering, an inflorescence architectural change, a change in stem bifurcations, a lack of a shoot meristem, reduced meristem cell differentiation, altered phyllotaxy, altered branching pattern, reduced apical dominance, reduced trichome density, ectopic trichome development, altered trichome development, altered stem morphology, increased root growth, increased root hairs, altered seed development, increased cell proliferation/cell differentiation, premature senescence, delayed senescence, lethality, increased necrosis, an increase in seedling or plant size, decreased plant size, a change in leaf morphology, increased altered leaf development, increased leaf size and mass, glossy leaves, leaf cell expansion, change in seed morphology, altered seed coloration, increased seed size, decreased seed size, altered seed shape, change in leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid content, increased leaf insoluble sugars, decreased leaf insoluble sugars, increased leaf anthocyanins, an alteration of leaf fatty acid content, an alteration of leaf glucosinolate content, change in seed biochemistry, an increase in seed oil content, decrease in seed oil content, increase in seed fatty acid content, decrease in seed fatty acid content, increase in seed protein content, decrease in seed protein content, alteration in seed prenyl lipid content, increase in seed sterols, upregulation of genes involved in secondary metabolism, increase in root anthocyanins, increase in plant anthocyanins, and alteration in light response or shade avoidance. The present sequence encodes an orthologue of a thalecress transcription factor isolated from Corn.

Sequence 1278 BP; 212 A; 409 C; 418 G; 239 T; 0 U; 0 Other;

Query Match 61.9%; Score 19.2; DB 12; Length 1278;
Best Local Similarity 87.5%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 TCACGGCATCGTCAGTGGCGCT 29
Db 577 TCACGGCGTCGTCAGTGGCGCT 554

RESULT 17

ID AD142860/C

XX AD142860 standard; DNA; 1278 BP.

AC AD142860;

XX 16-DEC-2004 (first entry)

DE Plant transcription factor polynucleotide #854.

transgenic; plant; enhanced tolerance to abiotic stress;
XX phosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;

KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
transcription factor; gene; da.

OS Zea mays.

XX US2004019927-A1.

PN 29-JAN-2004.

XX 25-FEB-2003; 2003US-00374780.

XX 18-APR-2001; 2001US-00837944.

PA (SHER/) SHERMAN B K.

PA (RIEC/) RIECHMANN J L.

PA (JIANG/) JIANG C.

PA (HEAR/) HEARD J E.

PA (HAAK/) HAAKE V.

PA (CREE/) CREELMAN R A.

PA (RATC/) RATCLIFFE O.

PA (ADAM/) ADAM L J.

PA (REUB/) REUBER T L.

PA (KEDD/) KEDDIE J.

PA (BROU/) BROUN P E.

PA (PILG/) PILGRIM M L.

PA (DUBE/) DUBELL A N.

PA (PINE/) PINEDA O.

PA (YUGG/) YU G.

XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;

PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;

PI Pilgrim ML, Dubell AN, Pineda O, Yu G;

XX WPI; 2004-132245/13.

XX New transgenic plant comprising a recombinant polynucleotide of any one

PT of more than 500 nucleotide sequences, useful in bioinformatic search

PT methods.

XX Claim 1; SEQ ID NO 1323; 435pp; English.

PS The invention describes a transgenic plant comprising a recombinant

CC polynucleotide of any one of more than 500 nucleotide sequences fully

CC defined in the specification or its complement. The method of the

CC invention can be used to produce a plant having altered traits such as:

CC enhanced tolerance to abiotic stress; glyphosate tolerance; hormone

CC sensitivity; disease resistance; sugar sensing; early or late flowering;

CC altered flower structure, change in stem bifurcations, altered branching

CC pattern, reduced apical dominance, reduced trichome density; lack of

CC trichomes; reduced ectopic trichome development; altered trichome

CC development; increase in trichome number; altered stem morphology;

CC increased root growth; increased root hairs; altered seed development;

CC altered cell proliferation or cell differentiation; rapid development;

CC premature senescence; increased necrosis; increase in seedling or plant

CC size; decreased plant size; leaf morphology; seed morphology; seed

CC biochemistry; increase in root anthocyanins; increase in plant

CC anthocyanins; or alteration in light response or shade avoidance. The

CC transgenic plant, polynucleotides and polypeptides are useful in

CC bioinformatic search methods. This sequence represents a plant

CC transcription factor, and an orthologue of Arabidopsis thaliana

CC transcription factors isolated in the invention, that can be used in the

CC creation of a transgenic plant with altered traits.

XX Sequence 1278 BP; 212 A; 409 C; 418 G; 239 T; 0 U; 0 Other;

XX Query Match 61.9%; Score 19.2; DB 13; Length 1278;

XX Best Local Similarity 87.5%; Pred. No. 2.8e+02;

XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX 6 TCACGGCATCGTCAGTGGCGCT 29

CC proliferation of an organism. The antiseptic nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1686 BP; 282 A; 464 C; 495 G; 445 T; 0 U; 0 Other;
Query Match 61.9%; Score 19.2; DB 8; Length 1686;
Best Local Similarity 87.5%; Pred. No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTCGGC 28
DB 785 CTCACGGCAGCATCAGTTCGGC 762
RESULT 20
ADL03792/c
ID ADL03792 standard; DNA; 1770 BP.
XX
AC ADL03792;
XX
DT 06-MAY-2004 (first entry)
XX
DE DNA encoding a M. catarrhalis protein #1478.
XX
KW ds; gene; Moraxella catarrhalis; infection.
XX
OS Moraxella catarrhalis.
XX
PN US6673910-B1.
XX
PD 06-JAN-2004.
XX
PF 04-APR-2000; 2000US-00540236.
XX
PR 08-APR-1999; 99US-0128416P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2004-178127/17.
DR P-PSDB; ADL05712.
XX
PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.
XX
PS Disclosure; SEQ ID NO 1478; 429pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
CC catarrhalis protein.
XX
SQ Sequence 1770 BP; 494 A; 400 C; 417 G; 459 T; 0 U; 0 Other;
Query Match 61.3%; Score 19; DB 12; Length 1770;
Best Local Similarity 81.5%; Pred. No. 3.6e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTCGGCTTG 31
DB 648 CTCACGGCATATCGGTTTCGGTTG 622

CC proliferation of an organism. The antiseptic nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1686 BP; 282 A; 464 C; 495 G; 445 T; 0 U; 0 Other;
Query Match 61.9%; Score 19.2; DB 8; Length 1686;
Best Local Similarity 87.5%; Pred. No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTCGGC 28
DB 785 CTCACGGCAGCATCAGTTCGGC 762
RESULT 20
ADL03792/c
ID ADL03792 standard; DNA; 1770 BP.
XX
AC ADL03792;
XX
DT 06-MAY-2004 (first entry)
XX
DE DNA encoding a M. catarrhalis protein #1478.
XX
KW ds; gene; Moraxella catarrhalis; infection.
XX
OS Moraxella catarrhalis.
XX
PN US6673910-B1.
XX
PD 06-JAN-2004.
XX
PF 04-APR-2000; 2000US-00540236.
XX
PR 08-APR-1999; 99US-0128416P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2004-178127/17.
DR P-PSDB; ADL05712.
XX
PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.
XX
PS Disclosure; SEQ ID NO 1478; 429pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
CC catarrhalis protein.
XX
SQ Sequence 1770 BP; 494 A; 400 C; 417 G; 459 T; 0 U; 0 Other;
Query Match 61.3%; Score 19; DB 12; Length 1770;
Best Local Similarity 81.5%; Pred. No. 3.6e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTCGGCTTG 31
DB 648 CTCACGGCATATCGGTTTCGGTTG 622

RESULT 21
AAF28530/c
ID AAF28530 standard; DNA; 23210 BP.
XX
AC AAF28530;
XX
DT 04-APR-2001 (first entry)
XX
DE Genomic fragment #17.
XX
KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW bronchopulmonary; endocarditis; meningitis; ss.
XX
OS Moraxella catarrhalis.
XX
PN WO200078968-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US016649.
XX
PR 18-JUN-1999; 99US-0140121P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lagace RE, Patterson C, Berg KL;
XX WPI; 2001-041427/05.
XX
PT Genomic library for identifying diagnostic and therapeutic compositions,
PT and for identifying virulence factors, regulatory elements and drug
PT targets, comprises Moraxella catarrhalis nucleic acids.
XX
PS Claim 1; Page 141-146; 545pp; English.
XX
CC The present invention relates to a Moraxella catarrhalis genomic library
CC comprising of a combination of 41 nucleic acid molecules (see AAF28514-
CC AAF28554). The library has a number of uses described in the
CC specification e.g. is useful for identifying diagnostic and therapeutic
CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC aerobic, gram-negative diplococcus, normally found among the bacterial
CC flora of human upper airways. M. catarrhalis is known to cause acute,
CC localised infections such as otitis media, sinusitis and bronchopulmonary
CC infection and life-threatening, systemic diseases including endocarditis
CC and meningitis
XX
SQ Sequence 23210 BP; 7108 A; 4482 C; 4951 G; 6668 T; 0 U; 1 Other;
Query Match 61.3%; Score 19; DB 4; Length 23210;
Best Local Similarity 81.5%; Pred. No. 5.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTCGGCTTG 31
DB 16504 CTCACGGCATATCGGTTTCGGTTG 16478
RESULT 22
AAI66165/c
ID AAI66165 standard; DNA; 41599 BP.
XX
AC AAI66165;
XX
DT 15-JAN-2002 (first entry)
XX
DE Bacillus subtilis Iturin A.
XX
KW Bacillus subtilis; Iturin A; cation channel; ds.
XX
OS Bacillus subtilis.
XX
PN JP2001231561-A.

```
XX 28-AUG-2001.
XX
XX PD
XX PF 18-FEB-2000; 2000JP-00040825.
XX
XX PR 18-FEB-2000; 2000JP-00040825.
XX
XX PA (MASA/) MASADA M.
XX
XX DR WPI; 2001-642167/74.
XX
XX PT A gene encoding Iturin A for the production of large amounts of Iturin A.
XX
XX PS Disclosure; Page 4-17; 18pp; Japanese.
XX
XX CC The invention relates to a gene encoding Iturin A encoding a protein
XX having an activity of promoting the transfer of a cation to the exterior
XX of the cell through the cation channel of the cell and a promoter for
XX transferring the cation to the exterior of the cell containing Iturin A
XX as the active component. The gene can be used for the preparation of
XX Iturin A in a large quantity
XX
XX SQ Sequence 41599 BP; 12054 A; 9173 C; 10241 G; 10131 T; 0 U; 0 Other;
XX
XX Query Match 61.3%; Score 19; DB 4; Length 41599;
XX Best Local Similarity 81.5%; Pred. No. 5.7e+02;
XX Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 5 CTCACGGCATCGTCAGTTGCGGCTTG 31
DB 4696 CTTAACGGCATCGTCATGTGCTGCTTG 4670

RESULT 23
ACH94680
ID ACH94680 standard; DNA; 1275 BP.
XX
XX AC ACH94680;
XX
XX DT 29-JUL-2004 (first entry)
XX
XX DE Klebsiella pneumoniae polynucleotide seqid 475.
XX
XX KW Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
XX OS Klebsiella pneumoniae.
XX
XX PN US6610836-B1.
XX
XX PD 26-AUG-2003.
XX
XX PF 27-JAN-2000; 2000US-00489039.
XX
XX PR 29-JAN-1999; 99US-0117747P.
XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX PI Breton GL, Osborne M;
XX
XX DR WPI; 2003-895346/82.
XX
XX DR P-PSDB; ABO61129.
XX
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX PS Disclosure; SEQ ID NO 475; 932pp; English.
XX
XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition

CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention
XX
XX SQ Sequence 1275 BP; 199 A; 360 C; 399 G; 317 T; 0 U; 0 Other;
XX
XX Query Match 60.6%; Score 18.8; DB 11; Length 1275;
XX Best Local Similarity 76.7%; Pred. No. 4.2e+02;
XX Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AGCACTCAACGGCATCGTCAGTTGCGGCTT 30
DB 393 AGCGCTCATTTGGATCGTCATCGCGGCTT 422

RESULT 24
ABL23899/C
ID ABL23899 standard; DNA; 2571 BP.
XX
XX AC ABL23899;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23170.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX PR 23-MAR-2000; 2000US-0191837P.
XX
XX PR 11-JUL-2000; 2000US-00614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX DR WPI; 2001-656860/75.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX PS Claim 1; SEQ ID NO 23170; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 2571 BP; 551 A; 765 C; 726 G; 529 T; 0 U; 0 Other;
XX
XX Query Match 60.6%; Score 18.8; DB 4; Length 2571;
XX Best Local Similarity 76.7%; Pred. No. 4.6e+02;
XX Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GCATCTCAACGGCATCGTCAGTTGCGGCTTG 31
DB 1043 GCATTGAGCGCATCTTTCAGTTTCCGCTCG 1014

RESULT 25
```

ABL23898
ID ABL23898 standard; DNA; 4571 BP.
XX
AC ABL23898;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23167.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 23167; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4571 BP; 1023 A; 1117 C; 1158 G; 1273 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 4; Length 4571;
Best Local Similarity 76.7%; Pred. No. 5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 GCACTCAACGCGATCGTCAGTTCGCGCTTG 31
|||||
Db 2529 GCATTGAGCGGCATCTTCAGTTCCGCTCG 2558
Search completed: March 11, 2005, 04:19:58
Job time : 47.3508 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 12.557 Seconds
(without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-13
Perfect score: 31
Sequence: 1 agacatcaaggcatcgatcggttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20.6	66.5	219	4	US-09-248-796A-7462
C 2	20.6	66.5	645	4	Sequence 7462, Ap
C 3	19	61.3	582	4	Sequence 7463, Ap
C 4	19	61.3	584	4	Sequence 9428, Ap
C 5	19	61.3	1770	4	Sequence 1817, Ap
C 6	19	61.3	23210	4	Sequence 1478, Ap
C 7	18.8	60.6	1275	4	Sequence 17, Appl
C 8	18.8	60.6	1500	4	Sequence 475, App
C 9	18.4	59.4	552	4	Sequence 12895, A
C 10	18.4	59.4	576	4	Sequence 12895, A
C 11	18.4	59.4	741	4	Sequence 1715, A
C 12	18.4	59.4	1011	4	Sequence 1715, A
C 13	18.4	59.4	1107	4	Sequence 1912, Ap
C 14	18.4	59.4	1410	4	Sequence 4174, Ap
C 15	18.4	59.4	22301	4	Sequence 16449, A
C 16	18.2	58.7	678	4	Sequence 16339, A
C 17	18.2	58.7	678	4	Sequence 1208, Ap
C 18	18.2	58.7	1500	4	Sequence 2734, Ap
C 19	18.2	58.7	1500	4	Sequence 18016, A
C 20	18.2	58.7	276687	4	Sequence 233, App
C 21	18	58.1	450	4	Sequence 13840, A
C 22	18	58.1	1058	4	Sequence 13840, A
C 23	18	58.1	1058	4	Sequence 654, App
C 24	18	58.1	1155	4	Sequence 639, App
C 25	18	58.1	1373	6	Sequence 6152, Ap
C 26	18	58.1	1728	6	Sequence 298, App
C 27	18	58.1	4403765	3	Patent No. 5514590-3
					Patent No. 5514590-3
					Sequence 616, App
					Sequence 2, Appl

28 58.1 4411529 3 US-09-103-840A-1 Sequence 1, Appli
29 17.8 57.4 285 4 US-09-270-767-191 Sequence 191, App
30 17.8 57.4 285 4 US-09-270-767-15473 Sequence 15473, A
31 17.8 57.4 452 4 US-09-582-200A-7 Sequence 7, Appli
32 17.8 57.4 601 4 US-09-949-016-191075 Sequence 191075,
33 17.8 57.4 1146 3 US-09-134-001C-906 Sequence 906, App
34 17.8 57.4 1317 4 US-09-252-991A-11745 Sequence 11745, A
35 17.8 57.4 1422 4 US-09-724-623-42 Sequence 42, Appl
36 17.8 57.4 1442 4 US-09-634-238-82 Sequence 82, Appl
37 17.8 57.4 1692 4 US-09-252-991A-11887 Sequence 11887, A
38 17.8 57.4 1867 1 US-07-955-905A-1 Sequence 1, Appli
39 17.8 57.4 3144 4 US-09-252-991A-12554 Sequence 12554, A
40 17.8 57.4 4431 4 US-09-252-991A-12856 Sequence 12856, A
41 17.8 57.4 24333 4 US-09-639-207-9 Sequence 9, Appli
42 17.8 57.4 26843 4 US-09-949-016-17208 Sequence 17208, A
43 17.8 57.4 169998 3 US-09-676-610B-24 Sequence 24, Appl
44 17.8 57.4 197496 4 US-09-877-177A-10 Sequence 10, Appl
45 17.6 56.8 601 4 US-09-949-016-28238 Sequence 28238, A

ALIGNMENTS

RESULT 1
US-09-248-796A-7462/C
; Sequence 7462, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7462
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7462

Query Match 66.5%; Score 20.6; DB 4; Length 219;
Best Local Similarity 85.2%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 ACTCAACGGCATCGTCAGTTGCGGCTT 30
Db 70 AGTCAACGGCATCGACAGTTGCGGCTT 44

RESULT 2
US-09-248-796A-7463
; Sequence 7463, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7463
; LENGTH: 645
; TYPE: DNA

```
; ORGANISM: Candida albicans
US-09-248-796A-7463

Query Match      66.5%; Score 20.6; DB 4; Length 645;
Best Local Similarity 85.2%; Pred. No. 8.5;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ACTCAACGGCATCGTCAGTTGCGGCTT 30
Db 149 AGTCAACGCCATCGACAGTTGCGTCTT 175

RESULT 3
US-09-540-236-1478/c
; Sequence 1478, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1478
; LENGTH: 1770
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-1478

Query Match      61.3%; Score 19; DB 4; Length 1770;
Best Local Similarity 81.5%; Pred. No. 58;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CTCAACGGCATCGTCAGTTGCGGCTTG 31
Db 648 CTCAACGGCATATCGGTTTCGGTTTG 622

RESULT 6
US-09-596-002-17/c
; Sequence 17, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 23210
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 28
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 17
; PUBLICATION INFORMATION:
US-09-596-002-17

Query Match      61.3%; Score 19; DB 4; Length 23210;
Best Local Similarity 81.5%; Pred. No. 1e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CTCAACGGCATCGTCAGTTGCGGCTTG 31
Db 16504 CTCAACGGCATATCGGTTTCGGTTTG 16478

RESULT 7
US-09-489-039A-475
; Sequence 475, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
```

```
; ORGANISM: Candida albicans
US-09-248-796A-7463

Query Match      66.5%; Score 20.6; DB 4; Length 645;
Best Local Similarity 85.2%; Pred. No. 8.5;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ACTCAACGGCATCGTCAGTTGCGGCTT 30
Db 149 AGTCAACGCCATCGACAGTTGCGTCTT 175

RESULT 3
US-09-540-236-1478/c
; Sequence 1478, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1478
; LENGTH: 1770
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-1478

Query Match      61.3%; Score 19; DB 4; Length 1770;
Best Local Similarity 81.5%; Pred. No. 58;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CTCAACGGCATCGTCAGTTGCGGCTTG 31
Db 648 CTCAACGGCATATCGGTTTCGGTTTG 622

RESULT 6
US-09-596-002-17/c
; Sequence 17, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 23210
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 28
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 17
; PUBLICATION INFORMATION:
US-09-596-002-17

Query Match      61.3%; Score 19; DB 4; Length 23210;
Best Local Similarity 81.5%; Pred. No. 1e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CTCAACGGCATCGTCAGTTGCGGCTTG 31
Db 16504 CTCAACGGCATATCGGTTTCGGTTTG 16478

RESULT 7
US-09-489-039A-475
; Sequence 475, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
```



```

US-09-270-767-18016/c
; Sequence 18016, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18016
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-18016

Query Match      58.7%; Score 18.2; DB 4; Length 678;
Best Local Similarity 74.2%; Pred.No.1.1e+00;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1  AGCACTCAACGGCATCGTCAGTTGCGGCTTG 31
      ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db      32  AGCTCTCAACGACATGGACAGCTTTTCACCTTG 2

RESULT 18
US-09-603-208A-233
; Sequence 233, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 233
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:

```

```
Query Match          58.1%; Score 18; DB 4; Length 450;
Best Local Similarity 80.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGCACTCAACGGCATCGTCAGTTGCG 26
DB 113 ACGGCTCAACGGCATCGCTGATGAG 138

RESULT 21
US-09-134-000C-639/c
; Sequence 639, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 639
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-639

Query Match          58.1%; Score 18; DB 4; Length 1008;
Best Local Similarity 80.8%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ACTCAACGGCATCGTCAGTTGGGCT 29
DB 316 ATTCAACGTCATCTTCAGTTGCAGCT 291

RESULT 22
US-09-902-540-6152
; Sequence 6152, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6152
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-6152

Query Match          58.1%; Score 18; DB 4; Length 1068;
Best Local Similarity 80.8%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACTCAACGGCATCGTCAGTTGCGGC 28
DB 158 CACTCAACGGCATCGCTACTGCTGC 183

RESULT 23
US-09-902-540-298/c
```

```
; Sequence 298, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 298
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1155)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-298

Query Match          58.1%; Score 18; DB 4; Length 1155;
Best Local Similarity 80.8%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACTCAACGGCATCGTCAGTTGCGGC 28
DB 912 CACTCAACGGCATCGGCTACTGCG 887

RESULT 24
5514590-3
; Patent No. 5514590
; APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.; JAMES, ERIC
; TITLE OF INVENTION: EXPRESSION SYSTEM COMPRISING DNA
; ENCODING THE SIGNAL PEPTIDE OF PROTEASE B FROM STREPTOMYCES
; GRISEUS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,644
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 66,938
; FILING DATE: 25-MAY-1993
; APPLICATION NUMBER: 844,937
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: 221,346
; FILING DATE: 18-JUL-1988
; APPLICATION NUMBER: 795,331
; FILING DATE: 06-NOV-1985
; SEQ ID NO:3;
; LENGTH: 1373
5514590-3

Query Match          58.1%; Score 18; DB 6; Length 1373;
Best Local Similarity 80.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GCATCAACGGCATCGTCAGTTGCGG 27
DB 1022 GCATCAACGCCACCGTCACTACGG 1047

RESULT 25
5514590-3
; Patent No. 5514590
; APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.; JAMES, ERIC
; TITLE OF INVENTION: EXPRESSION SYSTEM COMPRISING DNA
; ENCODING THE SIGNAL PEPTIDE OF PROTEASE B FROM STREPTOMYCES
; GRISEUS
```

```

; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,644
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 66,938
; FILING DATE: 25-MAY-1993
; APPLICATION NUMBER: 844,937
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: 221,346
; FILING DATE: 18-JUL-1988
; APPLICATION NUMBER: 795,331
; FILING DATE: 06-NOV-1985
; SEQ ID NO:3
; LENGTH: 1373
5514590-3

Query Match      58.1%; Score 18; DB 6; Length 1373;
Best Local Similarity 80.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      2 GCCTCAACGGCAGTCGTCAGTTGCGG 27
Db      1022 GCCTCAACGGCAGTCGTCAGTTGCGG 1047

```

Search completed: March 11, 2005, 13:13:09
Job time : 16.557 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	20.2	55.2	664	17	US-10-338-110-22	Sequence 22, Appl	
C 2	20.2	65.2	1156	17	US-10-282-122A-36496	Sequence 36496, A	
C 3	20.2	65.2	1662	17	US-10-282-122A-38840	Sequence 38840, A	
C 4	20	64.5	1632	17	US-10-369-493-35320	Sequence 35320, A	
C 5	19.8	63.9	434	17	US-10-424-599-6242	Sequence 6242, Ap	
C 6	19.4	62.6	96597	17	US-10-085-117-112	Sequence 112, App	
C 7	19.2	61.9	1247	17	US-10-425-114-35365	Sequence 35365, A	
C 8	19.2	61.9	1252	18	US-10-425-115-86277	Sequence 86277, A	
C 9	19.2	61.9	1278	17	US-10-374-780A-1323	Sequence 1323, Ap	
C 10	19.2	61.9	1278	17	US-10-412-699B-1472	Sequence 1472, Ap	
C 11	19.2	61.9	1362	17	US-10-369-493-44549	Sequence 44549, A	

```
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 36496
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (151)..(152)
; OTHER INFORMATION: n=g, a, t or c
US-10-282-122A-36496

Query Match 65.2%; Score 20.2; DB 17; Length 1156;
Best Local Similarity 88.0%; Pred. No. 55;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ACTCAACGGCATCTCAGTTGCGGC 28
Db 294 ACTCCAGGCAGCATCAGTTGCGGC 270
|||||
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 36496
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-10-282-122A-38840

Query Match 65.2%; Score 20.2; DB 17; Length 1662;
Best Local Similarity 88.0%; Pred. No. 57;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ACTCAACGGCATCTCAGTTGCGGC 28
Db 762 ACTCCAGGCAGCATCAGTTGCGGC 738
|||||
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 38840
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35920

Query Match 64.5%; Score 20; DB 17; Length 1632;
Best Local Similarity 82.1%; Pred. No. 69;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ACTCAACGGCATCTCAGTTGCGGCTTG 31
```

```
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 38840
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-10-282-122A-38840

Query Match 65.2%; Score 20.2; DB 17; Length 1662;
Best Local Similarity 88.0%; Pred. No. 57;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ACTCAACGGCATCTCAGTTGCGGC 28
Db 762 ACTCCAGGCAGCATCAGTTGCGGC 738
|||||
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35920
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35920

Query Match 64.5%; Score 20; DB 17; Length 1632;
Best Local Similarity 82.1%; Pred. No. 69;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ACTCAACGGCATCTCAGTTGCGGCTTG 31
```



```
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1323
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G987
US-10-374-780A-1323

Query Match 61.9%; Score 19.2; DB 17; Length 1278;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 TCACGGCATCGTCAGTGGCGGT 29
Db 577 TCACGGCGTCGTCAGTGGCGGT 554

RESULT 10
US-10-412-699B-1472/c
; Sequence 1472, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: Dubell, Arnold N.
; APPLICANT: Ratcliffe, Oliver.
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
```

```
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1472
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Zea mays
US-10-412-699B-1472

Query Match 61.9%; Score 19.2; DB 17; Length 1278;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 TCACGGCATCGTCAGTGGCGGT 29
Db 577 TCACGGCGTCGTCAGTGGCGGT 554

RESULT 11
US-10-369-493-44549
; Sequence 44549, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44549
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-44549

Query Match 61.9%; Score 19.2; DB 17; Length 1362;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 CTCACGGCATCGTCAGTGGCGC 28
Db 412 CGCAACGGCATCGTCATTGGGAC 435
```


RESULT 12
US-10-282-122A-39500/c
; Sequence 39500, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 39500
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Salmonella typhi
US-10-282-122A-39500

Query Match 61.9%; Score 19.2; DB 17; Length 1686;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTGCGGC 28
DB 785 CTCACGGCATCGTCAGTTGCGGC 762

RESULT 13
US-10-672-787-17/c
; Sequence 17, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELITRA.025C1
; CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 09/596,002
; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 23210
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 28
; OTHER INFORMATION: a or g or c or t, unknown, or other
US-10-672-787-17
Query Match 61.3%; Score 19; DB 17; Length 23210;
Best Local Similarity 81.5%; Pred. No. 2.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 CTCACGGCATCGTCAGTTGCGGCTG 31
DB 16504 CTCACGGCATATCGGTTTCGGTTG 16478
RESULT 14
US-10-712-672-3048/c
; Sequence 3048, Application US/10712672
; Publication No. US20040102413A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Chowrira, Bharat
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; TITLE OF INVENTION: Method and Reagent for the Inhibition of Telomerase Enzyme
; FILE REFERENCE: MHB00-882-C (400/019)
; CURRENT APPLICATION NUMBER: US/10/712,672
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US/09/653,225
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/197,769
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/150,713
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 5586
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3048
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-10-712-672-3048

Query Match 60.6%; Score 18.8; DB 18; Length 38;
Best Local Similarity 90.9%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 8 AACGGCATCGTCAGTTGCGGCT 29
DB 22 AACGGCATCATCAGTTGCGGCT 1

RESULT 15
US-10-437-963-96785
; Sequence 96785, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 96785
; LENGTH: 7017
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; LOCATION: (1)...(31477)
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9484C.1
US-10-437-963-96785

Query Match 60.6%; Score 18.8; DB 18; Length 7017;
Best Local Similarity 76.7%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GCACCTCAACGGCATCGTCAGTTGGCGCTTG 31
|||||
Db 6790 GCACCTCAACGGCTTGGACGGGTGTGGCATG 6819

RESULT 16
US-10-437-963-96807/c
; Sequence 96807, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 96807
; LENGTH: 7284
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9486C.1
US-10-437-963-96807

Query Match 60.6%; Score 18.8; DB 18; Length 7284;
Best Local Similarity 76.7%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GCACCTCAACGGCATCGTCAGTTGGCGCTTG 31
|||||
Db 228 GCACCTCAACGGCTTGGACGGGTGTGGCATG 199

RESULT 17
US-10-087-192-373/c
; Sequence 373, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 373
; LENGTH: 31477
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(31477)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-373

Query Match 60.6%; Score 18.8; DB 13; Length 31477;
Best Local Similarity 76.7%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AGACCTCAACGGCATCGTCAGTTGCGGCTT 30
|||||
Db 11683 AGACCTCAACGGCATGCCAGTTGAAGCTT 11654

RESULT 18
US-09-974-300-287/c
; Sequence 287, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-287

Query Match 60.0%; Score 18.6; DB 9; Length 917;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CTCACCGGCATCGTCAGTTGCGGCT 29
|||||
Db 324 CTCACCGGCAGCGCAATTGCGGCT 300

RESULT 19
US-10-369-493-25159/c
; Sequence 25159, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

```
; SEQ ID NO 25159
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-25159

Query Match          60.0%; Score 18.6; DB 17; Length 1056;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 TCAACGGCATCGTCAGTTGCGGCTT 30
Db 416 TGAACGGCATCGGCAGTTCGTATT 392

RESULT 20
US-09-815-242-5975/c
; Sequence 5975, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5975
; LENGTH: 3150
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3150)
US-09-815-242-5975

Query Match          60.0%; Score 18.6; DB 9; Length 3150;
Best Local Similarity 84.0%; Pred. No. 3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 CAACGGCATCGTCAGTTGCGGCTTG 31
Db 351 CAACGGCATCGCCAGCTGCAGTTTG 327

RESULT 21
US-10-282-122A-6959/c
; Sequence 6959, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

Query Match          60.0%; Score 18.6; DB 9; Length 3150;
Best Local Similarity 84.0%; Pred. No. 3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 CAACGGCATCGTCAGTTGCGGCTTG 31
Db 351 CAACGGCATCGCCAGCTGCAGTTTG 327

RESULT 22
US-10-297-465A-1
; Sequence 1, Application US/10297465A
; Publication No. US20040142413A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Andrew
; APPLICANT: Reinach, Fernando
; APPLICANT: Setubal, Joao
; APPLICANT: Medanis, Joao
; APPLICANT: Arruda, Paulo
; TITLE OF INVENTION: Isolated Genome of Xylella Fastidiosa and Uses Thereof
; FILE REFERENCE: PAFESP 202 US (10213376)
; CURRENT APPLICATION NUMBER: US/10/297,465A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01618
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,906
; PRIOR FILING DATE: 2001-06-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.2
```


Db 790 GCGCGATCGCACGGTCAGTTGCGGCT 817

Search completed: March 12, 2005, 00:25:20
Job time : 158.869 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 410.213 Seconds
(without alignment)
2876.537 Million cell updates/sec

Title: US-09-674-277-13
Perfect score: 31
Sequence: 1 agcactcaaggcagtcgctgagcggttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gsl1.*

9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21.4	69.0	653	8	AZ566286 219PVF07
C 2	21	67.7	491	6	CB176223 pl19e06.y
C 3	21	67.7	500	6	CA341518 pk22e12.x
C 4	21	67.7	504	6	CB276633 pl34e02.y
C 5	21	67.7	526	6	CB176022 pl17a03.y
C 6	21	67.7	544	6	CB174873 pl14a06.y
C 7	21	67.7	620	6	CB190567 pl28c08.y
C 8	20.6	66.5	319	7	CO269310 EK094106.
C 9	20.6	66.5	391	7	CO293017 EK201520.
C 10	20.6	66.5	515	1	AI107484 GH05184.5
C 11	20.6	66.5	523	7	CO340131 EP10059.3
C 12	20.6	66.5	583	1	AI135590 GH13404.5
C 13	20.2	65.2	271	8	AZ214463 Sheared D
C 14	20.2	65.2	414	8	BH254127 SALK 0160
C 15	20.2	65.2	548	9	TA176A11Q
C 16	20	64.5	337	9	BX214909 Danio rer
C 17	20	64.5	494	6	CB887967 GL18.106L
C 18	20	64.5	505	4	B1313636 dai29a06.
C 19	20	64.5	522	8	AZ048062 LMAJFV1.1
C 20	20	64.5	528	5	BQ400890 NISC mp11
C 21	20	64.5	571	5	BQ399395 NISC mp03
C 22	20	64.5	588	5	BQ399724 NISC mp05
C 23	20	64.5	593	4	B1441697 dai38h12.
C 24	20	64.5	629	9	BX242611 Danio rer

C 25	20	64.5	630	4	BG022425 dab03g03.
C 26	20	64.5	734	7	CF283935 AGENCOURT
C 27	20	64.5	798	7	CF342190 AGENCOURT
C 28	20	64.5	800	7	CF284555 AGENCOURT
C 29	20	64.5	818	6	CB560134 AGENCOURT
C 30	20	64.5	850	7	CF284554 AGENCOURT
C 31	20	64.5	854	6	CB565651 AGENCOURT
C 32	20	64.5	858	7	CF284556 AGENCOURT
C 33	20	64.5	860	6	CB562031 AGENCOURT
C 34	20	64.5	866	6	CB565253 AGENCOURT
C 35	20	64.5	876	6	CB562993 AGENCOURT
C 36	20	64.5	882	6	CB561012 AGENCOURT
C 37	20	64.5	892	9	CG824759 SOYAW62TH
C 38	20	64.5	931	6	CB560171 AGENCOURT
C 39	19.8	63.9	385	7	CK391871 K0836G10-
C 40	19.8	63.9	440	1	AU070098 AU070098
C 41	19.8	63.9	656	2	BB093704 BB093704
C 42	19.8	63.9	689	6	CA582795 EST002470
C 43	19.8	63.9	749	7	CN174587 AGENCOURT
C 44	19.8	63.9	847	4	B1110178 602899082
C 45	19.8	63.9	975	5	BUS16442 AGENCOURT

ALIGNMENTS

RESULT 1
AZ566286/c
LOCUS
DEFINITION 219PVF07 Pv MBN #16 (amplified twice) Plasmodium vivax genomic 3',
653 bp DNA linear GSS 07-MAY-2001
ACCESSION AZ566286
VERSION AZ566286.1 GI:13973225
KEYWORDS GSS.
SOURCE Plasmodium vivax (malaria parasite P. vivax)
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 653)
AUTHORS Carlton,J.M.-R. and Dame,J.B.
TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
COMMENT Contact: Dame JB
Dept. of Pathobiology, College of Veterinary Medicine
University of Florida
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
Tel: 352 392 4700
Fax: 352 392 9704
Email: damej@mail.vetmed.ufl.edu
Seq primer: M13(-20) forward
Class: shotgun.

FEATURES
source

Location/Qualifiers
1..653
/organism="Plasmodium vivax"
/mol_type="genomic DNA"
/strain="Belem"
/db_xref="taxon:5855"
/dev_stage="asexual blood forms"
/lab_hosts="Saimiri boliviensis"
/clone_lib="Pv MBN #16 (amplified twice)"
/notes="Vector: Lambda ZAP II (Stratagene); individual clones excised into phagemid plasmids; Site 1: Ecor I; Site 2: Ecor I; Genomic DNA was prepared from asynchronous blood stage forms of the Belem line of P. vivax grown in squirrel monkeys. Parasitized erythrocytes were purified from contaminating host leukocytes by filtration of ADP activated blood through acid-washed glass beads and Whatman CF11 cellulose columns by gravity filtration. Purified DNA was digested with mung bean nuclease in the presence of 42.5% formamide at 50°C as described (Galinski, M. et al. 1992. Cell 69,1213-1226; Vernick, K.D. et al. 1988. N.A.R. 16, 6883-6896). Eco RI linkers were added and the constructs ligated into Lambda ZAP II. P. vivax Belem was originally isolated from a patient in

RESULT 9	CO293017/c	391 bp	mRNA	linear	EST 25-JUN-2004
LOCUS	CO293017				
DEFINITION	Exelixa FlyTag CK01 pCDNA-SK+ Drosophila melanogaster cDNA clone EK201520 5, mRNA sequence.				
ACCESSION	CO293017				
KEYWORDS	CO293017.1	GI:49214563			
SOURCE	EST.				
ORGANISM	Drosophila melanogaster (fruit fly)				
REFERENCE	Drosophila melanogaster				
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
TITLE	1 (bases 1 to 391)				
JOURNAL	Kopczynski, C., Platt, D., Campbell, J., Muzong, C., Laufer, A., Peterson, E. and Swimmer, C.				
COMMENT	Exelixa FlyTag EST Project CK01 Library Unpublished (2004)				
	Contact: Stapleton, M. BDGP				
	Lawrence Berkeley National Lab				
	One Cyclotron Rd, Berkeley, CA 94720, USA				
	Fax: 510 486 6798				
	Email: http://www.fruitfly.org/EST , est@fruitfly.berkeley.edu				
	Plate: EK.2015 row: B column: 8				
	High quality sequence stop: 279.				
FEATURES	Location/Qualifiers				
source	1..391				
	/organism="Drosophila melanogaster"				
	/mol_type="mRNA"				
	/db_xref="taxon:7227"				
	/clone="EK201520"				
	/note="Organ: mixed stage embryos, imaginal disks, and adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2: XhoI; Random primed, normalized library from mixed stage embryos, imaginal disks, and adult heads."				
ORIGIN					
Query Match	66.5%;	Score 20.6;	DB 7;	Length 391;	
Best Local Similarity	85.2%;	Pred. No. 2.6e+02;			
Matches	23;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
Qy	1 AGCACTCAACGGCATCGTCAGTTGCGG 27				
Db	177 AGCCCTTAATGGCATCATCAGTTGCGG 151				
RESULT 10	AI107484/c	515 bp	mRNA	linear	EST 02-DEC-2003
LOCUS	AI107484				
DEFINITION	GH051844.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH05184 5 similar to ferrochelatase: FBan002098 'enzyme' located on: 3R 100E3-100E3; 04/13/2001, mRNA sequence.				
ACCESSION	AI107484				
VERSION	AI107484.1	GI:3475137			
KEYWORDS	EST.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 515)				
TITLE	Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Teang, G., Lewis, S. and Rubin, G.M.				
JOURNAL	BDGP/HMI Drosophila EST Project				
COMMENT	Unpublished (2001)				
	Contact: Stapleton, M. BDGP				
	Lawrence Berkeley National Lab				
	One Cyclotron Rd, Berkeley, CA 94720, USA				

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGCACTCAACGGCATCGTCAGTTCGG 27
 Db 451 AGCCCTTAATGGCATCATCAGTTCGG 477

RESULT 12
 AII135590/c
 LOCUS
 DEFINITION
 Sheared DNA-104B7.TF Sheared DNA Trypanosoma brucei genomic clone
 Sheared DNA-104B7, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.

REFERENCE
 1 (bases 1 to 271)
 El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Sub, E., Malek, J., Fujii, C.,

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGCACTCAACGGCATCGTCAGTTCGG 27
 Db 152 AGCCCTTAATGGCATCATCAGTTCGG 126

RESULT 13
 A2214463
 LOCUS
 DEFINITION
 Sheared DNA-104B7.TF Sheared DNA Trypanosoma brucei genomic clone
 Sheared DNA-104B7, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.

REFERENCE
 1 (bases 1 to 271)
 El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Sub, E., Malek, J., Fujii, C.,

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGCACTCAACGGCATCGTCAGTTCGG 27
 Db 152 AGCCCTTAATGGCATCATCAGTTCGG 126

RESULT 13
 A2214463
 LOCUS
 DEFINITION
 Sheared DNA-104B7.TF Sheared DNA Trypanosoma brucei genomic clone
 Sheared DNA-104B7, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.

REFERENCE
 1 (bases 1 to 271)
 El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Sub, E., Malek, J., Fujii, C.,

Gerrard, C., Leesch, V., de Jong, P., Ullu, E., Melville, S.,
 Donelson, J., Fraser, C. and Adams, M.
 Determination of clone end sequences from Trypanosoma brucei GUTat
 10.1 sheared DNA library
 Unpublished (1999)
 Contact: Najib M. El-Sayed
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: nelsayed@tigr.org
 Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
 DNA library constructed at TIGR. Clones will be available for
 distribution through Research Genetics, Alabama, USA. Sheared DNA
 end sequences search page: <http://www.tigr.org/cdb/mdb/tbdb/>.
 Seq primer: M13-Forward
 Class: shotgun.

FEATURES
 Location/Qualifiers
 1..271
 /organism="Trypanosoma brucei"
 /mol_type="genomic DNA"
 /strain="TREU927/4 GUTat 10.1"
 /db_xref="taxon:5691"
 /clone="Sheared DNA-104B7"
 /clone_lib="Sheared DNA"
 /note="Vector: pUC18; Site 1: SmaI; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from a cloned population of
 Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
 sheared to give a tight size distribution (approx 2 kb).
 The v + i method used for the library construction is
 described in detail in Smith, H.O. and Venter, J.C.
 (Making small insert libraries for whole genome shotgun
 sequencing projects. In Genome Sequencing: A Practical
 Approach, eds. M. Vaudin and B. Borell, Oxford University
 Press, 1999)."

ORIGIN
 Query Match 65.2%; Score 20.2; DB 8; Length 271;
 Best Local Similarity 88.0%; Pred. No. 3.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 CAACGGCATCGTCAGTTCGGCTTG 31
 Db 214 CTACGGCATCGTCGGTTCGGCTTG 238

RESULT 14
 BH254127
 LOCUS
 DEFINITION
 SALK_016043 Arabidopsis thaliana TDNA insertion lines Arabidopsis
 thaliana genomic clone SALK_016043, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 414)
 Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
 Gadrinab, C., Jenke, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
 Shinn, P., Zimmerman, J. and Ecker, J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)
 Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752

Fax: 858 558 6379
Email: ecker@saik.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
At3g61340.
Class: TDNA tagged.
Location/Qualifiers
1. 414
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_016043"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.saik.edu/tdna_protocols.html"

FEATURES
source
Query Match 65.2%; Score 20.2; DB 8; Length 414;
Best Local Similarity 88.0%; Pred. No. 4e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

ORIGIN
6 TCACGGCATCGTCAGTTGGCGTT 30
|||||
Db 2 TCACGGCATCGTCATTTGCTTCT 26
|||||

RESULT 15
TA176A11Q/c
LOCUS
DEFINITION T. brucei sheared genomic DNA clone 176a11, reverse sequence,
genomic survey sequence.
ACCESSION AL475531
VERSION AL475531.1 GI:11840315
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 548)
Hall, N., Bowman, S., Lennard, N. J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
Melville, S. E., Rajandream, M. A. and Barrell, B. G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nh@sanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 CUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J. C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nh@sanger.ac.uk
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.
Location/Qualifiers
1. 548
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="176a11"

FEATURES
source
Query Match 64.5%; Score 20; DB 9; Length 337;
Best Local Similarity 82.1%; Pred. No. 4.8e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

ORIGIN
3 CACTCACGGCATCGTCAGTTGGCGTT 30
|||||
Db 152 CCCTCACAGCATGTGTCAGTTGGGGTT 179
|||||

RESULT 17
CD887967
LOCUS
DEFINITION G118.106110F010606 G118 Triticum aestivum cDNA clone G118106L10,
mRNA sequence.
ACCESSION CD887967
VERSION CD887967.1 GI:32654802
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 494)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french

Query Match 65.2%; Score 20.2; DB 9; Length 548;
Best Local Similarity 88.0%; Pred. No. 4.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 CAACGGCATCGTCAGTTGGCGTTG 31
|||||
Db 247 CTACGGCATGTCGGTTGGCGTTG 223
|||||

RESULT 16
BX214909
LOCUS
DEFINITION Danio rerio genomic clone DKEY-256F17, genomic survey sequence.
ACCESSION BX214909
VERSION BX214909.1 GI:28046795
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 337)
Humphray, S. J., Huckle, E. and Durham, J. L.
Direct Submission
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 256F17. 256F17
is part of the Daniokey BAC Library created by R. Plasterk and N. V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.
Location/Qualifiers
1. 337
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-256F17"
/tissue_type="Testis"
/note="vector pindigoBAC-536"

FEATURES
source
Query Match 82.1%; Score 20; DB 9; Length 337;
Best Local Similarity 82.1%; Pred. No. 4.8e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

ORIGIN
3 CACTCACGGCATCGTCAGTTGGCGTT 30
|||||
Db 152 CCCTCACAGCATGTGTCAGTTGGGGTT 179
|||||

RESULT 17
CD887967
LOCUS
DEFINITION G118.106110F010606 G118 Triticum aestivum cDNA clone G118106L10,
mRNA sequence.
ACCESSION CD887967
VERSION CD887967.1 GI:32654802
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 494)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french

plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
and <http://genoplante-info.infobiogen.fr>.

FEATURES

source
1. .494
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G118106L10"
/tissue_type="grain (118 degrees per day after
pollination)"
/clone_lib="G118"

ORIGIN

Query Match 64.5%; Score 20; DB 6; Length 494;
Best Local Similarity 82.1%; Pred. No. 5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CACTCAACGCGCATCGTCAGTTGCGGCTT 30

Db 173 CACTCAACAGCAGCAGCAGTGTGCTT 200

RESULT 18

BI313636/c
LOCUS
DEFINITION
dai29a06.x1 NICHDXGC Sp1 Xenopus laevis cDNA clone IMAGE:4963090
3' similar to FR:Q9W7Q0 Q9W7Q0 ELASTASE 3 PRECURSOR. ;, mRNA
sequence.

ACCESSION
BI313636

VERSION
BI313636.1 GI:14987963

KEYWORDS
EST.

SOURCE
Xenopus laevis (African clawed frog)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE
1 (bases 1 to 505)

AUTHORS
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL
Unpublished (1997)

COMMENT
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: Xenopus clones from this library are available

through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 380.

FEATURES

source
Location/Qualifiers
1. .505
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4963090"
/dev_stage="adult"
/clone_lib="NICHDXGC Sp1"
/lab_host="DH10B (phage-resistant)"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.4 kb. Constructed by Life
Technologies."

ORIGIN

Query Match 64.5%; Score 20; DB 4; Length 505;
Best Local Similarity 82.1%; Pred. No. 5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GCACTCAACGCGATCGTCAGTTGCGGCTT 29

Db 173 GCACTCAACGCGATCGTCAGTTGCGGCTT 29

Db 262 GAAGTCCACGCGCATTCAGTTTCGGCT 235

RESULT 19

AZ048062

LOCUS

DEFINITION

Leishmania major genomic clone LMAJFV1_lm72c06 3', genomic survey
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

1 (bases 1 to 522)

A survey of the Leishmania major Friedlin strain V1 genome by

shotgun sequencing: a resource for DNA microarrays and expression

profiling

Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)

11292569

11292569

Contact: Akopyants, NS / Beverley, SM

WashU Leishmania Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library construction: Natalia S. Akopyants, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

If using this information please cite:

N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major

Friedlin strain V1 genome by shotgun sequencing' and the Washington

University Genome Sequencing Center For information on obtaining

clone material please contact: Natalia S. Akopyants Ph.D.

(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.

(Beverley@borcim.wustl.edu)

Seq primer: -40UP from Gibco

Class: shotgun

High quality sequence stop: 452.

Location/Qualifiers

1. .522

/organism="Leishmania major"

/mol_type="genomic DNA"

/strain="Friedlin strain V1"

/db_xref="taxon:5664"

/clone="LMAJFV1_lm72c06"

/lab_host="TOP10 (Invitrogen)"

/clone_lib="Leishmania major FV1 random genomic library"

/notes="Vector: pZero-2 (Invitrogen); Site 1: EcoRV;

Genomic DNA was isolated from stationary phase cells. For

this library, DNA was sheared to give a tight size

distribution of 1-1.5kb fragments, blunt-ended with T4 DNA

polymerase, dephosphorylated with Shrimp Alkaline

Phosphatase and ligated into pZero-2 vector's EcoRV site."

ORIGIN

Query Match 64.5%; Score 20; DB 8; Length 522;

Best Local Similarity 82.1%; Pred. No. 5e+02;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ACTCAACGCGCATCGTCAGTTGCGGCTTG 31

Db 22 ACGCAACGCGCGTCGCTGCTTGCCTCG 49

RESULT 20

BQ400890/c

LOCUS

BQ400890

528 bp

linear

EST 22-MAY-2002

```

DEFINITION NISC mp11h07.x1 NICHDXGC_Sp1 Xenopus laevis cDNA clone
IMAGE:4969212 3', mRNA sequence.
ACCESSION BQ400890
VERSION BQ400890.1 GI:21088577
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus;
1 (bases 1 to 528)
NIH-XGC http://image.llnl.gov/image/html/xenopuslib.info.shtml.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D. )
Email: cgapbs-remail.nih.gov
cDNA Library Preparation:
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM10950 row: P column: 13
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..528
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4969212"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC_Sp1"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.4 kb. Constructed by Life
Technologies."

FEATURES
source
ORIGIN
Query Match 64.5%; Score 20; DB 5; Length 528;
Best Local Similarity 82.1%; Pred. No. 5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GCATCAACGCGCATGTCAGTTGCGGCT 29
| | | | | | | | | | | | | | | | | | | | |
Db 291 GAAGTCCACGCGCATGTCAGTTGCGGCT 264

RESULT 21
BQ399395/c
LOCUS
DEFINITION NISC mp03c12.x1 NICHDXGC_Sp1 Xenopus laevis cDNA clone
IMAGE:4965526 3', mRNA sequence.
ACCESSION BQ399395
VERSION BQ399395.1 GI:21087082
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus;
1 (bases 1 to 571)
NIH-XGC http://image.llnl.gov/image/html/xenopuslib.info.shtml.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
cDNA Library Preparation:
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM10949 row: C column: 17
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..586
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4965526"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC_Sp1"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.4 kb. Constructed by Life
Technologies."

ORIGIN
Query Match 64.5%; Score 20; DB 5; Length 528;
Best Local Similarity 82.1%; Pred. No. 5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GCATCAACGCGCATGTCAGTTGCGGCT 29
| | | | | | | | | | | | | | | | | | | | |
Db 291 GAAGTCCACGCGCATGTCAGTTGCGGCT 264

RESULT 21
BQ399395/c
LOCUS
DEFINITION NISC mp05b09.x1 NICHDXGC_Sp1 Xenopus laevis cDNA clone
IMAGE:4968520 3', mRNA sequence.
ACCESSION BQ399724
VERSION BQ399724.1 GI:21087411
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus;
1 (bases 1 to 588)
NIH-XGC http://image.llnl.gov/image/html/xenopuslib.info.shtml.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
cDNA Library Preparation:
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM10949 row: C column: 17
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..588
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4968520"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC_Sp1"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.4 kb. Constructed by Life
Technologies."

ORIGIN

```

```

Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM10941 row: F column: 23
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..571
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4965526"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC_Sp1"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.4 kb. Constructed by Life
Technologies."

ORIGIN
Query Match 64.5%; Score 20; DB 5; Length 571;
Best Local Similarity 82.1%; Pred. No. 5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GCATCAACGCGCATGTCAGTTGCGGCT 29
| | | | | | | | | | | | | | | | | | | | |
Db 244 GAAGTCCACGCGCATGTCAGTTGCGGCT 217

RESULT 22
BQ399724/c
LOCUS
DEFINITION NISC mp05b09.x1 NICHDXGC_Sp1 Xenopus laevis cDNA clone
IMAGE:4968520 3', mRNA sequence.
ACCESSION BQ399724
VERSION BQ399724.1 GI:21087411
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus;
1 (bases 1 to 588)
NIH-XGC http://image.llnl.gov/image/html/xenopuslib.info.shtml.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
cDNA Library Preparation:
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM10949 row: C column: 17
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..588
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4968520"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC_Sp1"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.4 kb. Constructed by Life
Technologies."

ORIGIN

```

```

Query Match      64.5%; Score 20; DB 5; Length 588;
Best Local Similarity 82.1%; Pred. No. 5.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GCATCAACGGCATCGTCAGTTGGGCT 29
    |||||
Db 256 GAAGTCCACGGCATGTGTCAGTTGGGCT 229
    |||||

RESULT 23
BI441697/c
LOCUS      BI441697      593 bp      mRNA      linear      EST 21-AUG-2001
DEFINITION dai38h12.x1 NICH.D_XGC_Sp1 Xenopus laevis cDNA clone IMAGE:4964207
3' similar to TR:Q9W7Q0 Q9W7Q0 ELASTASE 3 PRECURSOR. ;, mRNA
sequence.
ACCESSION   BI441697
VERSION     BI441697.1 GI:15266387
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM    Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 593)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgaps-rc@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            High quality sequence stop: 463.

FEATURES             source
    1..593
    /organism="Xenopus laevis"
    /mol_type="mRNA"
    /db_xref="taxon:8355"
    /clone="IMAGE:4964207"
    /dev_stage="adult"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NICH.D_XGC_Sp1"
    /note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: NotI;
    Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert size 2.4 kb. Constructed by Life
    Technologies."

ORIGIN
Query Match      64.5%; Score 20; DB 4; Length 593;
Best Local Similarity 82.1%; Pred. No. 5.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GCATCAACGGCATCGTCAGTTGGGCT 29
    |||||
Db 304 GAAGTCCACGGCATGTGTCAGTTACGGCT 277
    |||||

RESULT 24
BX242611
LOCUS      BX242611      629 bp      DNA      linear      GSS 29-JAN-2003
DEFINITION Danio rerio genomic clone DKEY-246M24, genomic survey sequence.
ACCESSION   BX242611
VERSION     BX242611.1 GI:28164945
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

```

```

Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 629)
Humphray,S.J., Huckle,E. and Durham,J.L.
Direct Submission
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Unpublished
This sequence was generated from the SP6 end of BAC 246M24. 246M24
is part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES             source
    1..629
    /organism="Danio rerio"
    /mol_type="genomic DNA"
    /db_xref="taxon:7955"
    /clone="DKEY-246M24"
    /tissue_type="Testis"
    /note="vector pindigoBAC-536"

ORIGIN
Query Match      64.5%; Score 20; DB 9; Length 629;
Best Local Similarity 82.1%; Pred. No. 5.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACTCAACGGCATCGTCAGTTGGGCTT 30
    |||||
Db 148 CCTCAACAGCATGTGTCAGTTGGGGTT 175
    |||||

RESULT 25
BG022425/c
LOCUS      BG022425      630 bp      mRNA      linear      EST 30-JAN-2001
DEFINITION dab03g03.x1 NICH.D_XGC_Sp1 Xenopus laevis cDNA clone IMAGE:4173485
3' similar to TR:Q9W7Q0 Q9W7Q0 ELASTASE 3 PRECURSOR. ;, mRNA
sequence.
ACCESSION   BG022425
VERSION     BG022425.1 GI:12478504
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM    Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 630)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Other ESTs: dab03g03.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgaps-rc@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 414.

FEATURES             source
    1..630
    /organism="Xenopus laevis"
    /mol_type="mRNA"
    /db_xref="taxon:8355"
    /clone="IMAGE:4173485"
    /dev_stage="adult"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NICH.D_XGC_Sp1"
    /note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: NotI;
    Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert size 2.4 kb. Constructed by Life
    Technologies."

```


ORIGIN

Query Match 64.5%; Score 20; DB 4; Length 630;
 Best Local Similarity 82.1%; Pred. No. 5.1e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GCACCTCAACGGCATCGTCAGTTGCGGCT 29
 Db 261 GAAGTCCACGGCATGTCAGTTGCGGCT 234

Search completed: March 11, 2005, 13:01:12
 Job time : 412.213 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 20.481 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-18

Perfect score: 14

Sequence: 1 ggcacgtcagttg 14

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: Geneseqn16Dec04:*
2: Geneseqn1980s:*
3: Geneseqn1990s:*
4: Geneseqn2000s:*
5: Geneseqn2001as:*
6: Geneseqn2001bs:*
7: Geneseqn2002as:*
8: Geneseqn2002bs:*
9: Geneseqn2003as:*
10: Geneseqn2003bs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	100.0	14	3	Aaz36118 Primer de
2	14	100.0	16	3	Aaz36119 Primer de
3	14	100.0	18	3	Aaz36120 Primer de
4	14	100.0	31	3	Aaz36112 Primer de
5	14	100.0	31	3	Aaz36113 Primer de
6	14	100.0	319	10	Adk56969 Plant DNA
7	14	100.0	319	10	Adk55125 Plant DNA
8	14	100.0	340	10	Adk53158 Plant DNA
9	14	100.0	368	10	Adk56429 Plant DNA
10	14	100.0	774	5	Aah65885 C glutami
11	14	100.0	825	8	ACA00240 C. glutam
12	14	100.0	1285	8	ACA49060 Prokaryot
13	14	100.0	1489	3	Aaz36101 Nucleic a
14	14	100.0	1551	8	ACA52263 Prokaryot
15	14	100.0	5145	4	Aaf24991 Nucleotid
16	14	100.0	349980	5	Aah65826 C glutami
17	13	92.9	60	6	ABN42551 Human epl
18	13	92.9	60	13	AD53708 Eucalyptu
19	13	92.9	132	4	Aai21101 Probe #11
20	13	92.9	132	4	ABA66178 Human foe

C 21	13	92.9	132	4	AAI46363 Probe #15
C 22	13	92.9	132	4	ABA48294 Human bre
C 23	13	92.9	132	4	ABA33245 Probe #11
C 24	13	92.9	132	4	AAK40343 Human bon
C 25	13	92.9	132	4	AAK14598 Human bra
C 26	13	92.9	132	4	AAK14598 Human bra
C 27	13	92.9	132	4	ABE33914 Human liv
C 28	13	92.9	132	5	AAI06820 Probe #68
C 29	13	92.9	132	6	ABE14361 Human gen
C 30	13	92.9	380	8	ACA35036 Prokaryot
C 31	13	92.9	397	9	ACH30566 Human tes
C 32	13	92.9	448	4	AAI11893 Probe #18
C 33	13	92.9	448	4	ABE33593 Human foe
C 34	13	92.9	448	4	AAI33220 Probe #19
C 35	13	92.9	448	4	ABA43174 Human bre
C 36	13	92.9	448	4	ABA23348 Probe #18
C 37	13	92.9	448	4	AAK01860 Human bon
C 38	13	92.9	448	4	ABE26893 Human liv
C 39	13	92.9	448	5	AAI01830 Probe #18
C 40	13	92.9	448	6	ABE01839 Human gen
C 41	13	92.9	464	3	AAF10758 Fusarium
C 42	13	92.9	550	6	ABQ92852 Triticum
C 43	13	92.9	589	13	ACN49193 Cotton pr
C 44	13	92.9	831	2	AAQ06129 Sequence
C 45	13	92.9	885	5	AAS79900 DNA encod

ALIGNMENTS

RESULT 1

AAZ36118
ID AAZ36118 standard; DNA; 14 BP.

XX AAZ36118;

DT 11-FEB-2000 (first entry)

DE Primer derived from a nucleic acid sequence specific to EHEC.

XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
XX PCR primer; probe; ss.
OS Synthetic.
OS Escherichia coli.
XX WO9955908-A2.
XX PD 04-NOV-1999.
XX PF 27-APR-1999; 99WO-FR001000.
XX PR 28-APR-1998; 98FR-00005329.
XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX Frechon DTM, Laure FC, Thierry D;
XX WPI; 2000-013443/01.
XX New nucleic acid containing sequences specific to enterohemorrhagic
XX Escherichia coli, particularly serotype O157:H7, used for detecting these
XX bacteria in food.
XX Claim 5; Page 27; 48pp; French.
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
XX specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
XX derived from two sequences. The first (AAZ36101) is 99.9% homologous to
XX the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
XX 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
XX The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 14 BP; 2 A; 3 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 3; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATCGTCAGTTG 14
 |||||
 Db 1 GGCATCGTCAGTTG 14

RESULT 2
 AAZ36119
 ID AAZ36119 standard; DNA; 16 BP.
 XX AC AAZ36119;
 XX DT 11-FEB-2000 (first entry)
 DE Primer derived from a nucleic acid sequence specific to EHEC.
 XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 KW PCR primer; probe; ss.
 XX Synthetic.
 OS Escherichia coli.
 XX WO9955908-A2.
 XX PD 04-NOV-1999.
 XX PF 27-APR-1999; 99WO-FR001000.
 XX PR 28-APR-1998; 98FR-00005329.
 XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
 XX PI Frechon DTM, Laure FC, Thierry D;
 XX WPI; 2000-013443/01.
 XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.
 XX Claim 5; Page 27; 48pp; French.

AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 16 BP; 2 A; 5 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATCGTCAGTTG 14
 |||||
 Db 1 GGCATCGTCAGTTG 14

RESULT 4
 AAZ36112
 ID AAZ36112 standard; DNA; 31 BP.

Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATCGTCAGTTG 14
 |||||
 Db 2 GGCATCGTCAGTTG 15

RESULT 3
 AAZ36120
 ID AAZ36120 standard; DNA; 18 BP.
 XX AC AAZ36120;
 XX DT 11-FEB-2000 (first entry)
 DE Primer derived from a nucleic acid sequence specific to EHEC.
 XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 KW PCR primer; probe; ss.
 XX Synthetic.
 OS Escherichia coli.
 XX WO9955908-A2.
 XX PD 04-NOV-1999.
 XX PF 27-APR-1999; 99WO-FR001000.
 XX PR 28-APR-1998; 98FR-00005329.
 XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
 XX PI Frechon DTM, Laure FC, Thierry D;
 XX WPI; 2000-013443/01.
 XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.
 XX Claim 5; Page 27; 48pp; French.

AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 18 BP; 3 A; 5 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 3; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATCGTCAGTTG 14
 |||||
 Db 3 GGCATCGTCAGTTG 16

RESULT 4
 AAZ36112
 ID AAZ36112 standard; DNA; 31 BP.


```

PF 30-AUG-2002; 2002WO-US027884.
XX
PR 31-AUG-2001; 2001US-0316471P.
XX
XX (DOWC ) DOW CHEM CO.
PA (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX
XX WPI; 2003-313091/30.
DR
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 4352; 2576pp; English.
PS
XX
XX The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC altered glyceride metabolism, altered ester metabolism, altered
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX
XX Sequence 319 BP; 82 A; 77 C; 93 G; 67 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 14; DB 10; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
Db |||||
45 GGCATCGTCAGTTG 58

RESULT 7
ADK55125
ID ADK55125 standard; DNA; 319 BP.
XX
XX AC ADK55125;
XX
XX 06-MAY-2004 (first entry)
XX
XX Plant DNA sequence which confers altered metabolic characteristic #2508.
DE
XX altered metabolic characteristic; plant; acid metabolism;
XX alcohol metabolism; fatty acid metabolism;
XX branched fatty acid metabolism; alkaloid metabolism;
XX amino acid metabolism; ester metabolism; glyceride metabolism;
XX phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
XX Unidentified.
OS
XX
XX WO2003020936-A1.
PN
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
XX 31-AUG-2001; 2001US-0316471P.
XX
XX (DOWC ) DOW CHEM CO.
PA
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX
XX WPI; 2003-313091/30.
DR
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 4352; 2576pp; English.
PS
XX
XX The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC altered glyceride metabolism, altered ester metabolism, altered
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX
XX Sequence 319 BP; 82 A; 77 C; 93 G; 67 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 14; DB 10; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
Db |||||
45 GGCATCGTCAGTTG 58

RESULT 8
ADK53158
ID ADK53158 standard; DNA; 340 BP.
XX
XX AC ADK53158;
XX
XX 06-MAY-2004 (first entry)
XX
XX Plant DNA sequence which confers altered metabolic characteristic #541.
DE
XX altered metabolic characteristic; plant; acid metabolism;
XX alcohol metabolism; fatty acid metabolism;
XX branched fatty acid metabolism; alkaloid metabolism;
XX amino acid metabolism; ester metabolism; glyceride metabolism;
XX phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
XX Hypocrea lixii.
OS
XX
XX WO2003020936-A1.
PN
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
XX 31-AUG-2001; 2001US-0316471P.
XX
XX (DOWC ) DOW CHEM CO.
PA (DOWC ) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX
XX WPI; 2003-313091/30.
DR
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 2508; 2576pp; English.
PS
XX
XX The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC altered glyceride metabolism, altered ester metabolism, altered
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX
XX Sequence 319 BP; 82 A; 77 C; 93 G; 67 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 14; DB 10; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
Db |||||
45 GGCATCGTCAGTTG 58

```

DR WPI; 2003-313091/30.

XX Novel genes that confer altered metabolic characteristics in Nicotiana

PT benhamiana plants, useful for altering the levels of metabolites e.g.

PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.

XX

PS Claim 1; SEQ ID NO 541; 2576pp; English.

XX

CC The invention comprises DNA sequences which confer an altered metabolic

CC characteristic when they are expressed in a plant. The DNA sequences of

CC the invention are useful for producing plants with an altered metabolic

CC characteristic, such as: altered acid metabolism, alcohol metabolism,

CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other

CC base metabolism, altered amino acid metabolism, altered ester metabolism,

CC altered glyceride metabolism, altered phenolic metabolism, altered

CC carbohydrate metabolism, alkene or alkyne metabolism, hydrocarbon

CC metabolism, ketone or quinone metabolism. The DNA sequences of the

CC invention may be used to provide disease resistance in a plant and gene

CC shuffling or sexual PCR procedures. The present nucleic acid represents a

CC DNA sequence of the invention.

XX

SQ Sequence 340 BP; 89 A; 81 C; 102 G; 68 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 10; Length 340;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14

Db 68 GGCATCGTCAGTTG 81

RESULT 9

ID ADK56429 standard; DNA; 368 BP.

XX

AC ADK56429;

XX

XX 06-MAY-2004 (first entry)

DT

DE Plant DNA sequence which confers altered metabolic characteristic #3812.

XX

XX altered metabolic characteristic; plant; acid metabolism;

XX alcohol metabolism; fatty acid metabolism;

XX branched fatty acid metabolism; alkaloid metabolism;

XX amino acid metabolism; ester metabolism; glyceride metabolism;

XX phenolic metabolism; carbohydrate metabolism; sterol metabolism;

XX terpene metabolism; isoprenoid metabolism; alkene metabolism;

XX alkyne metabolism; hydrocarbon metabolism; ketone metabolism;

XX quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.

XX

OS Hypocrea lixii.

XX

XX WO2003020936-A1.

PN

XX 13-MAR-2003.

PD

XX 30-AUG-2002; 2002WO-US027884.

PF

XX 31-AUG-2001; 2001US-0316471P.

PR

XX (DOWC) DOW CHEM CO.

PA

XX (DOWC) DOW AGROSCIENCES LLC.

PA

XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;

PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;

PI

XX WPI; 2003-313091/30.

DR

XX Novel genes that confer altered metabolic characteristics in Nicotiana

PT benhamiana plants, useful for altering the levels of metabolites e.g.

PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.

PT

XX Claim 1; SEQ ID NO 3812; 2576pp; English.

XX

CC The invention comprises DNA sequences which confer an altered metabolic

CC characteristic when they are expressed in a plant. The DNA sequences of

CC the invention are useful for producing plants with an altered metabolic

CC characteristic, such as: altered acid metabolism, alcohol metabolism,

CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other

CC base metabolism, altered amino acid metabolism, altered ester metabolism,

CC altered glyceride metabolism, altered phenolic metabolism, altered

CC carbohydrate metabolism, alkene or alkyne metabolism, hydrocarbon

CC metabolism, ketone or quinone metabolism. The DNA sequences of the

CC invention may be used to provide disease resistance in a plant and gene

CC shuffling or sexual PCR procedures. The present nucleic acid represents a

CC DNA sequence of the invention.

XX

SQ Sequence 368 BP; 117 A; 82 C; 101 G; 68 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 10; Length 368;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14

Db 68 GGCATCGTCAGTTG 81

RESULT 10

AAH65885/C

ID AAH65885 standard; DNA; 774 BP.

XX

AC AAH65885;

XX

XX 26-SEP-2001 (first entry)

DT

XX C glutamicum coding sequence fragment SEQ ID NO: 920.

DE

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis; ds.

KW

XX Corynebacterium glutamicum.

OS

XX EP1108790-A2.

PN

XX 20-JUN-2001.

PD

XX 18-DEC-2000; 2000EP-00127688.

PF

XX 16-DEC-1999; 99JP-00377484.

PR

XX 07-APR-2000; 2000JP-00159162.

PR

XX 03-AUG-2000; 2000JP-00280988.

PR

XX (KYOW) KYOWA HAKKO KOGYO KK.

PA

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

PI

XX WPI; 2001-376931/40.

DR

XX P-PSDB; AAG90665.

DR

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

PT

XX Claim 8; SEQ ID NO 920; 246pp + Sequence Listing; English.

XX

CC The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analysing

CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homologue of a gene derived from

CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX
 SQ Sequence 774 BP; 184 A; 202 C; 184 G; 204 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 5; Length 774;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
 |||||
 DB 420 GGCATCGTCAGTTG 407

RESULT 11

ACA00240/C
 ID ACA00240 standard; DNA; 825 BP.
 XX
 AC ACA00240;
 DT 02-JUN-2003 (first entry)
 XX C. glutamicum derived ORF SEQ ID 231.
 DE
 XX Coryneform; nucleic acid array; fermentation; culture; ds.
 KW Corynebacterium glutamicum.
 OS
 XX DB10128510-A1.
 PN
 XX 19-DEC-2002.
 PD
 XX 13-JUN-2001; 2001DE-01028510.
 PF
 XX 13-JUN-2001; 2001DE-01028510.
 PR
 XX (DEGS) DEGUSSA AG.
 PA
 XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
 PI
 XX WPI; 2003-279970/28.
 DR
 XX New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 XX
 PS Claim 1; Page 151-152; 709pp; German.
 XX
 CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX
 SQ Sequence 825 BP; 194 A; 214 C; 199 G; 216 T; 0 U; 2 Other;

Query Match 100.0%; Score 14; DB 8; Length 825;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
 |||||
 DB 471 GGCATCGTCAGTTG 458

RESULT 12

ACA49060
 ID ACA49060 standard; DNA; 1285 BP.
 XX
 AC ACA49060;
 DT 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #30717.
 DE
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Salmonella paratyphi.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 PF
 XX 21-MAR-2002; 2002WO-US009107.
 PR
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 XX WPI; 2003-029926/02.
 DR
 XX P-PSDB; ABU45190.
 DR
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX
 PS Claim 14; SEQ ID NO 36930; 1766pp; English.

XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1285 BP; 306 A; 347 C; 328 G; 304 T; 0 U; 0 Other;
 Query Match 100.0%; Score 14; DB 8; Length 1285;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAATCGTCAGTTG 14
 |||||
 Db 403 GGCAATCGTCAGTTG 416

RESULT 13
 AA236101
 ID AA236101 standard; DNA; 1489 BP.
 XX
 AC AA236101;
 XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.
 XX
 KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 KW I891; ds.
 XX
 OS Escherichia coli.
 XX
 PN WO9955908-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-FR001000.
 XX
 PR 28-APR-1998; 98FR-00005329.
 XX
 PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
 XX
 PI Frechon DTM, Laure FC, Thierry D;
 XX
 DR WPI; 2000-013443/01.
 XX
 PT New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.
 XX
 PS Claim 1; Fig 1; 48pp; French.
 XX
 CC The present sequence is specific to enterohemorrhagic Escherichia coli
 CC (EHEC). The sequence is 99.9% homologous to the katP gene of E. coli
 CC O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
 CC homologous with IS91 of E. coli (nucleotides 1-406 of the present
 CC sequence). The present sequence is of plasmid origin. Fragments of the
 CC present sequence are used, as probes and primers, for detection of E.
 CC coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
 CC animal samples, foods or the environment. The fragments are also useful
 CC for epidemiological studies
 XX
 SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;
 Query Match 100.0%; Score 14; DB 3; Length 1489;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAATCGTCAGTTG 14
 |||||
 Db 397 GGCAATCGTCAGTTG 410

RESULT 14
 ACA52263
 ID ACA52263 standard; DNA; 1551 BP.
 XX
 AC ACA52263;
 XX

19-JUN-2003 (first entry)
 Prokaryotic essential gene #33920.
 Antisense; ds; prokaryotic essential gene; cell proliferation;
 drug design; gene.
 Salmomella typhi.
 WO200277183-A2.
 03-OCT-2002.
 21-MAR-2002; 2002WO-US009107.
 21-MAR-2001; 2001US-00815242.
 06-SEP-2001; 2001US-00948993.
 25-OCT-2001; 2001US-0342923P.
 08-FEB-2002; 2002US-00072851.
 06-MAR-2002; 2002US-0362699P.
 (ELIT-) ELITRA PHARM INC.
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 WPI; 2003-029926/02.
 P-PSDB; ABU48393.
 New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.
 Claim 14; SEQ ID NO 40133; 1766pp; English.
 The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than S. aureus, S. typhimurium,
 K. pneumoniae or P. aeruginosa. The present sequence is one of the target
 prokaryotic essential genes. Note: The sequence data for this patent did
 not form part of the printed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

Sequence 1551 BP; 358 A; 441 C; 404 G; 348 T; 0 U; 0 Other;
 Query Match 100.0%; Score 14; DB 8; Length 1551;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
 DB 694 GGCATCGTCAGTTG 707

RESULT 15
 AAF24991/c
 ID AAF24991 standard; cDNA; 5145 BP.
 XX
 AC AAF24991;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Nucleotide sequence of a Drosophila polypeptide designated orbit.
 XX
 KW Orbit; chromosome segregation; mitotic spindle; microtubule; mitosis;
 KW polyploid cell; microtubule binding; nucleation; motor protein; cancer;
 KW leukaemia cell; solid tumour; ss.
 XX
 OS Drosophila sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 66..4544
 FT /*tag= a
 FT /product= "Orbit"
 XX
 XX W0200104295-A1.
 XX
 PD 18-JAN-2001.
 XX
 PF 11-JUL-2000; 2000WO-GB002662.
 XX
 PR 13-JUL-1999; 99GB-00016402.
 PR 24-DEC-1999; 99GB-00030707.
 XX
 PA (UYDU-) UNIV DUNDEE.
 XX
 PI Avides MDC, Deak P, Glover DM;
 XX
 DR WPI; 2001-138345/14.
 DR P-PSDB; AAB31597.
 XX
 PT Novel Drosophila orbit protein and its human homolog, useful for treating
 PT tumors and in assays for identifying substances capable of inhibiting
 PT mitosis.
 XX
 PS Claim 3; Page 84-85; 92pp; English.
 XX
 CC The present sequence encodes a Drosophila polypeptide designated orbit.
 CC The orbit protein is required for chromosome segregation. Orbit protein
 CC localises to mitotic spindles and binds microtubules. Orbit mutant
 CC phenotypes result in embryos exhibiting abnormal mitosis and polyploid
 CC cells. Orbit polypeptides are useful for inhibiting mitosis, and for
 CC identifying substances which affect orbit functions such as microtubule
 CC binding, microtubule organizing centre, nucleation activity and
 CC interactions with microtubule motor proteins. Tests for detecting or
 CC sequencing orbit, or its homologue, in a biological sample may be used to
 CC determine orbit sequences within cells in individuals who have or are
 CC suspected to have, an altered orbit gene sequence, for example within
 CC cancer cells including leukaemia cells and solid tumours such as breast,
 CC ovary, lung, colon etc
 XX
 SQ Sequence 5145 BP; 1388 A; 1331 C; 1295 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 4; Length 5145;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
 DB 3749 GGCATCGTCAGTTG 3736

RESULT 16
 AAH68526
 ID AAH68526 standard; DNA; 349980 BP.
 XX
 AC AAH68526;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 7061.
 XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Disclosure; SEQ ID NO 7061; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 349980 BP; 79274 A; 90638 C; 98727 G; 81341 T; 0 U; 0 Other;

Query Match 100.0%; Score 14; DB 5; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTTG 14
 DB 270272 GGCATCGTCAGTTG 270285

RESULT 17
 ABN42551/c
 ID ABN42551 standard; DNA; 60 BP.
 XX
 AC ABN42551;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:15299.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;


```

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX Claim 25; SEQ ID NO 11034; 487pp; English.
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging of
XX diseases of the cervix, notably cervical cancer. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 132 BP; 25 A; 46 C; 36 G; 25 T; 0 U; 0 Other;
XX
Query Match 92.9%; Score 13; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 GGCATCGTCAGTT 13
Db 15 GGCATCGTCAGTT 3
XX
RESULT 20
ABA66178/c
ID ABA66178 standard; DNA; 132 BP.
XX
AC ABA66178;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #14483.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
Qy 1 GGCATCGTCAGTT 13
Db 15 GGCATCGTCAGTT 3
XX
RESULT 21
AAI46363/c
ID AAI46363 standard; DNA; 132 BP.
XX
AC AAI46363;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #15049 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX Claim 25; SEQ ID NO 15049; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENPs).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders
XX

```

SQ Sequence 132 BP; 25 A; 46 C; 36 G; 25 T; 0 U; 0 Other;
Query Match 92.9%; Score 13; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCATCGTCAGTT 13
DB 15 GGCATCGTCAGTT 3

RESULT 22
ABA48294/c
ID ABA48294 standard; DNA; 132 BP.
XX
AC ABA48294;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #6989.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001;
XX
PF 30-JAN-2001; 2001WO-US0000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 4; SEQ ID NO 6989; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 25 A; 46 C; 36 G; 25 T; 0 U; 0 Other;

Query Match 92.9%; Score 13; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCATCGTCAGTT 13
DB 15 GGCATCGTCAGTT 3

RESULT 23
ABA33245/c
ID ABA33245 standard; DNA; 132 BP.
XX
AC ABA33245;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #11711 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 11711; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 25 A; 46 C; 36 G; 25 T; 0 U; 0 Other;

Query Match 92.9%; Score 13; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCATCGTCAGTT 13
DB 15 GGCATCGTCAGTT 3

```
RESULT 24
AAK40343/C
ID AAK40343 standard; DNA; 132 BP.
XX
AC AAK40343;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 14900.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234587P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 14900; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
SQ Sequence 132 BP; 25 A; 46 C; 36 G; 25 T; 0 U; 0 Other;
Query Match 92.9%; Score 13; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 15 GGCATCGTCAGTT 3
|||||
|||||

RESULT 25
AAK14598/C
ID AAK14598 standard; DNA; 132 BP.
XX
AC AAK14598;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 14589.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 14900; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
SQ Sequence 132 BP; 25 A; 46 C; 36 G; 25 T; 0 U; 0 Other;
Query Match 92.9%; Score 13; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATCGTCAGTT 13
Db 15 GGCATCGTCAGTT 3
|||||
|||||

Search completed: March 11, 2005, 04:20:00
Job time : 22.481 secs
```

Result No.	Score	Query Match	Length	DB	ID	Description	
C 1	13	92.9	425	4	US-09-270-767-13941	Sequence 13941, A	
C 2	13	92.9	601	4	US-09-949-016-108325	Sequence 108325,	
C 3	13	92.9	601	4	US-09-949-016-175637	Sequence 175637,	
C 4	13	92.9	864	4	US-09-949-016-4991	Sequence 4991, Ap	
C 5	13	92.9	1299	1	US-08-453-473-3	Sequence 3, Appli	
C 6	13	92.9	1299	1	US-08-038-948-3	Sequence 3, Appli	
C 7	13	92.9	1299	1	US-08-453-953-3	Sequence 3, Appli	
C 8	13	92.9	1299	2	US-08-862-903-3	Sequence 3, Appli	
C 9	13	92.9	1299	2	US-08-484-1588-60	Sequence 60, Appl	
C 10	13	92.9	1587	4	US-09-489-039A-6670	Sequence 6670, Ap	
C 11	13	92.9	2115	4	US-08-388-852B-1	Sequence 1, Appli	
C 12	13	92.9	2130	3	US-09-056-105-1	Sequence 1, Appli	
C 13	13	92.9	2131	4	US-08-234-784B-91	Sequence 91, Appli	
C 14	13	92.9	2154	1	US-08-448-170-3	Sequence 3, Appli	
C 15	13	92.9	2154	3	US-08-961-803-3	Sequence 3, Appli	
C 16	13	92.9	2172	2	US-08-417-174-26	Sequence 26, Appl	
C 17	13	92.9	2172	2	US-08-231-565A-26	Sequence 26, Appl	
C 18	13	92.9	2172	2	US-09-007-961-26	Sequence 26, Appl	
C 19	13	92.9	2172	3	US-09-267-439-26	Sequence 26, Appl	
C 20	13	92.9	2172	4	US-09-073-138-26	Sequence 26, Appl	
C 21	13	92.9	2328	4	US-09-252-991A-5729	Sequence 5729, Ap	
C 22	13	92.9	2469	4	US-09-489-039A-6749	Sequence 6749, Ap	
C 23	13	92.9	2568	4	US-09-328-352-2090	Sequence 2090, Ap	
C 24	13	92.9	2781	3	US-09-021-560-3	Sequence 3, Appli	
C 25	13	92.9	2781	4	US-09-202-178A-2	Sequence 2, Appli	
C 26	13	92.9	3340	3	US-09-021-560-1	Sequence 1, Appli	
C 27	13	92.9	3340	4	US-09-202-178A-1	Sequence 1, Appli	

US-09-949-016-108325

Query Match 92.9%; Score 13; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCATCGTCAGTTG 14
|||||
DB 352 GCATCGTCAGTTG 364

RESULT 3

US-09-949-016-175637/c
; Sequence 175637, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 175637
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-175637

Query Match 92.9%; Score 13; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCGTCAGTT 13
|||||
DB 380 GGATCGTCAGTT 368

RESULT 4

US-09-949-016-4991/c
; Sequence 4991, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4991
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4991

Query Match 92.9%; Score 13; DB 4; Length 864;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCGTCAGTT 13
|||||
DB 357 GGATCGTCAGTT 345

RESULT 5

US-08-453-472-3/c
; Sequence 3, Application US/08453472
; Patent No. 5626846
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; POLYPEPTIDES
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,472
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US3
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: ZP3
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP3 cdna
US-08-453-472-3

Query Match 92.9%; Score 13; DB 1; Length 1299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
| | | | | | | | | | | | | | |
Db 337 GGCATCGTCAGTT 325

RESULT 6

US-08-038-948-3/c
; Sequence 3, Application US/08038948
; Patent No. 5641487
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON
; TITLE OF INVENTION: ALLOIMMUNIZATION WITH ZONA PELLUCIDA POLYPEPTIDES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,948
; FILING DATE: 26-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: 99152/E-266-88/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-038-948-3

Query Match 92.9%; Score 13; DB 1; Length 1299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
| | | | | | | | | | | | | | |
Db 337 GGCATCGTCAGTT 325

RESULT 7

US-08-453-952-3/c
; Sequence 3, Application US/08453952
; Patent No. 5672488
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE

; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,952
; FILING DATE: 30-May-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: ZP3
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP3 cDNA
; US-08-453-952-3

Query Match 92.9%; Score 13; DB 1; Length 1299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
| | | | | | | | | | | | | | |
Db 337 GGCATCGTCAGTT 325

RESULT 8

US-08-862-903-3/c
; Sequence 3, Application US/08862903
; Patent No. 5916768

GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,903
; FILING DATE: 30-May-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: ZP3
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP3 cdna

US-08-862-903-3
Query Match 92.9%; Score 13; DB 2; Length 1299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTT 13
Db 337 GGCATCGTCAGTT 325
RESULT 9

US-08-484-158B-60/c
; Sequence 60, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immunocontraception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "human ZPC"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1282
; US-08-484-158B-60

Query Match 92.9%; Score 13; DB 2; Length 1299;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTT 13
Db 337 GGCATCGTCAGTT 325
RESULT 10
US-09-489-039A-6670/c
; Sequence 6670, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

```
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6670
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6670

Query Match          92.9%; Score 13; DB 4; Length 1587;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGCATCGTCAGTT 13
Db      810 GGCATCGTCAGTT 798

RESULT 11
US-08-388-852B-1/c
; Sequence 1, Application US/08388852B
; Patent No. 6500919
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan; Figdor, Carl Gustav.
; TITLE OF INVENTION: Melanoma associated antigenic polypeptide,
; TITLE OF INVENTION: epitopes thereof and vaccine against melanoma.
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Adema, Gosse Jan; Figdor, Carl Gustav
; STREET: Philips van Leydenlaan 25
; CITY: Nijmegen
; STATE: Brabant
; COUNTRY: the Netherlands
; ZIP: 6525 EX
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,852B
; FILING DATE: February 15, 1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Melanoma
; CELL TYPE: Melanocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22...2005
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 1...81
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1792...1870
; OTHER INFORMATION: /function = "transmembrane region"
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 262...264
; OTHER INFORMATION: /bound moiety = "carbohydrate"
; FEATURE:
```

```
; NAME/KEY: misc binding
; LOCATION: 337...339
; OTHER INFORMATION: /bound moiety = "carbohydrate"
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 352...354
; OTHER INFORMATION: /bound moiety = "carbohydrate"
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 982...984
; OTHER INFORMATION: /bound moiety = "carbohydrate"
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 1723...1725
; OTHER INFORMATION: /bound moiety = "carbohydrate"
US-08-388-852B-1

Query Match          92.9%; Score 13; DB 4; Length 2115;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGCATCGTCAGTT 13
Db      408 GGCATCGTCAGTT 396

RESULT 12
US-09-056-105-1/c
; Sequence 1, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; TITLE OF INVENTION: PROCESSING
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2130
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-1

Query Match          92.9%; Score 13; DB 3; Length 2130;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGCATCGTCAGTT 13
Db      408 GGCATCGTCAGTT 396

RESULT 13
US-08-234-784B-91/c
; Sequence 91, Application US/08234784B
; Patent No. 6660276
; GENERAL INFORMATION:
; APPLICANT: Slingluff, Craig L, Jr.
; APPLICANT: Engelhard, Victor H.
; APPLICANT: Hunt, Donald F.
; APPLICANT: Shabanowitz, Jeffrey
; APPLICANT: Cox, Andrea L.
; TITLE OF INVENTION: PEPTIDES RECOGNIZED BY MELANOMA-SPECIFIC
; TITLE OF INVENTION: CYTOTOXIC LYMPHOCYTES, AND USES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
```

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/234,784B
FILING DATE: 29-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/197,399
FILING DATE: 16-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: SLINGLUFF=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 2131
TYPE: nucleic acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: complete gene
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: melanocyte
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: chromosome 12
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: SEQ ID NO: 91
LOCATION: gene for Pmel-17 (accession number M77348)
IDENTIFICATION METHOD: This gene encodes the protein Pmel-17,
IDENTIFICATION METHOD: which was found to contain Seq ID No. 6660276 14, which h
IDENTIFICATION METHOD: ivity as an epitope for melanoma-specific cytotoxic T lym
OTHER INFORMATION: This sequence has previously been des-
OTHER INFORMATION: cribed. We are claiming its use in a tumor vaccine for the p
OTHER INFORMATION: and treatment of melanoma.
PUBLICATION INFORMATION:
AUTHORS: Kwon, B. S. et al.
TITLE: A Melanocyte Specific Gene, Pmel-17, Mapped Near the
TITLE: Silver Coat Color Locus on Mouse Chromosome 10 and is in A Syntenic
TITLE: Region on Human Chromosome 12
JOURNAL: Proc. Natl. Acad. Sci. USA
VOLUME: 88
ISSUE:
PAGES: 9228-9232
DATE: 1991
US-08-234-784B-91

Query Match 92.9%; Score 13; DB 4; Length 2131;
Best Local Similarity 100.0%; Pred. No. 3e+02;
; Sequence 3, Application US/08961803

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/234,784B
FILING DATE: 29-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/197,399
FILING DATE: 16-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: SLINGLUFF=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 2131
TYPE: nucleic acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: complete gene
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: melanocyte
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: chromosome 12
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: SEQ ID NO: 91
LOCATION: gene for Pmel-17 (accession number M77348)
IDENTIFICATION METHOD: This gene encodes the protein Pmel-17,
IDENTIFICATION METHOD: which was found to contain Seq ID No. 6660276 14, which h
IDENTIFICATION METHOD: ivity as an epitope for melanoma-specific cytotoxic T lym
OTHER INFORMATION: This sequence has previously been des-
OTHER INFORMATION: cribed. We are claiming its use in a tumor vaccine for the p
OTHER INFORMATION: and treatment of melanoma.
PUBLICATION INFORMATION:
AUTHORS: Kwon, B. S. et al.
TITLE: A Melanocyte Specific Gene, Pmel-17, Mapped Near the
TITLE: Silver Coat Color Locus on Mouse Chromosome 10 and is in A Syntenic
TITLE: Region on Human Chromosome 12
JOURNAL: Proc. Natl. Acad. Sci. USA
VOLUME: 88
ISSUE:
PAGES: 9228-9232
DATE: 1991
US-08-234-784B-91

Query Match 92.9%; Score 13; DB 1; Length 2154;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 GCATCGTCAGTT 13
Db 398 GCATCGTCAGTT 386

RESULT 14
US-08-448-170-3/C
; Sequence 3, Application US/08448170
; Patent No. 5723758
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 5723758el Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. PSI58C2, Active Against Lepidopteran Pests, and Genes
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,170
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 102D.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2154 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-448-170-3

Query Match 92.9%; Score 13; DB 1; Length 2154;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 2 GCATCGTCAGTTG 14
Db 351 GCATCGTCAGTTG 339

RESULT 15
US-08-961-803-3/c
; Sequence 3, Application US/08961803

Patent No. 6150589
GENERAL INFORMATION:
APPLICANT: Payne, Jewel
APPLICANT: Cummings, David A.
APPLICANT: Cannon, Raymond J.C.
APPLICANT: Narva, Kenneth E.
APPLICANT: Steiman, Steve
TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
TITLE OF INVENTION: B.t. P815822, Active Against Lepidopteran Pests, and Genes
TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jay M. Sanders
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,803
FILING DATE: 31-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/069,902
FILING DATE: 01-JUNE-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/759,247
FILING DATE: 13-SEPT-1991
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/448,170
FILING DATE: 23-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: M/S 102DCD1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2154 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-961-803-3
Query Match 92.9%; Score 13; DB 3; Length 2154;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 GCATCGTCAGTTG 14
DB 351 GCATCGTCAGTTG 339
RESULT 16
US-08-417-174-26/c
Sequence 26, Application US/08417174
Patent No. 5844075
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,565A
FILING DATE: 22-APR-1994

TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 126
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,174
FILING DATE: 05-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/231,565
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172
TYPE: nucleotide
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: CDNA
US-08-417-174-26
Query Match 92.9%; Score 13; DB 2; Length 2172;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCATCGTCAGTT 13
DB 425 GGCATCGTCAGTT 413
RESULT 17
US-08-231-565A-26/c
Sequence 26, Application US/08231565A
Patent No. 5874560
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,565A
FILING DATE: 22-APR-1994

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172
TYPE: nucleotide
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cdna
US-08-231-565A-26

Query Match 92.9%; Score 13; DB 2; Length 2172;
Best Local Similarity 100.0%; Pred. No. 3e+02; 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCGTCGTCAGTT 13
|||||
DB 425 GGCGTCGTCAGTT 413

RESULT 18
US-09-007-961-26/c
Sequence 26, Application US/09007961
Patent No. 5994523
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
METHODS
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/007,961
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/231,565
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172
TYPE: nucleotide
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cdna
US-09-007-961-26

Query Match 92.9%; Score 13; DB 2; Length 2172;
Best Local Similarity 100.0%; Pred. No. 3e+02; 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCGTCGTCAGTT 13
|||||
DB 425 GGCGTCGTCAGTT 413

RESULT 19
US-09-267-439-26/c
Sequence 26, Application US/09267439
Patent No. 6270778
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
METHODS
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 126
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/267,439
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/417,174
FILING DATE: 05-APR-1995
APPLICATION NUMBER: US/08/231,565
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172
TYPE: nucleotide
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cdna
US-09-267-439-26

Query Match 92.9%; Score 13; DB 3; Length 2172;
Best Local Similarity 100.0%; Pred. No. 3e+02; 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCGTCGTCAGTT 13
|||||
DB 425 GGCGTCGTCAGTT 413

RESULT 20
US-09-073-138-26/c
Sequence 26, Application US/09073138
Patent No. 6537560
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
METHODS
NUMBER OF SEQUENCES: 126
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,138
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/417,174
FILING DATE: 05-APR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172
TYPE: nucleotide
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: CDNA
US-09-073-138-26

Query Match 92.9%; Score 13; DB 4; Length 2172;
Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 425 GGCATCGTCAGTT 413

RESULT 21
US-09-252-991A-5729
; Sequence 5729, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5729
; LENGTH: 2328
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5729

Query Match 92.9%; Score 13; DB 4; Length 2328;

Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCATCGTCAGTT 13
|||||
Db 290 GGCATCGTCAGTT 302
RESULT 22
US-09-489-039A-6749
; Sequence 6749, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6749
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6749

Query Match 92.9%; Score 13; DB 4; Length 2469;
Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
|||||
Db 1699 GGCATCGTCAGTT 1711

RESULT 23
US-09-328-352-2090/c
; Sequence 2090, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2090
; LENGTH: 2568
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2090

Query Match 92.9%; Score 13; DB 4; Length 2568;
Best Local Similarity 100.0%; Pred. No. 3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GCATCGTCAGTTG 14
|||||
Db 1850 GCATCGTCAGTTG 1838

RESULT 24
US-09-021-560-3
; Sequence 3, Application US/09021560
; Patent No. 6410719
; GENERAL INFORMATION:
; APPLICANT: BOREN, THOMAS
; APPLICANT: NORMARK, STAFFAN
; APPLICANT: ARNOVIST, ANNA
; APPLICANT: ILVER, DAG

; TITLE OF INVENTION: BLOOD GROUP ANTIGEN BINDING PROTEIN AND
 ; TITLE OF INVENTION: CORRESPONDING GENES
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: BIRCH, STEWART, KOLASCH AND BIRCH
 ; STREET: PO BOX 747
 ; CITY: FALLS CHURCH
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/021,560
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURPHY JR, GERALD M
 ; REGISTRATION NUMBER: 28,977
 ; REFERENCE/DOCKET NUMBER: 825-144P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 205-8000
 ; TELEFAX: (703) 205-8050
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2781 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 409..2532
 ; US-09-021-560-3

Query Match 92.9%; Score 13; DB 3; Length 2781;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
 |||||
 Db 1963 GGCATCGTCAGTT 1975

RESULT 25
 US-09-202-178A-2
 ; Sequence 2, Application US/09202178A
 ; Patent No. 6709656
 ; GENERAL INFORMATION:
 ; APPLICANT: BOREN, Thomas
 ; APPLICANT: ARNOVIST, Anna
 ; APPLICANT: HAMMARSTROM, Lennart
 ; APPLICANT: NORMARK, Staffan
 ; APPLICANT: ILVER, Dag
 ; TITLE OF INVENTION: HELICOBACTER PYLORI ADHESIN BINDING GROUP ANTIGEN
 ; FILE REFERENCE: 0825-150P
 ; CURRENT APPLICATION NUMBER: US/09/202,178A
 ; CURRENT FILING DATE: 1999-02-10
 ; PRIOR APPLICATION NUMBER: PCT/SE97/01009
 ; PRIOR FILING DATE: 1997-06-10
 ; PRIOR APPLICATION NUMBER: 9602287-6 SE
 ; PRIOR FILING DATE: 1996-06-10
 ; PRIOR APPLICATION NUMBER: 9701014-4 SE
 ; PRIOR FILING DATE: 1997-03-19
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2781
 ; TYPE: DNA
 ; ORGANISM: Helicobacter pylori

US-09-202-178A-2

Query Match 92.9%; Score 13; DB 4; Length 2781;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTT 13
 |||||
 Db 1963 GGCATCGTCAGTT 1975

Search completed: March 11, 2005, 13:13:11
 Job time : 7.67089 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 185.257 Seconds
2876.537 Million cell updates/sec

Title: US-09-674-277-18
Perfect score: 14
Sequence: 1 ggcacgtcagtg 14

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14	100.0	204	1	AV269567
C 2	14	100.0	275	9	LBAP040807
C 3	14	100.0	323	9	AG241349
C 4	14	100.0	324	4	BM726445
C 5	14	100.0	438	6	CB486369
C 6	14	100.0	456	6	CA995113
C 7	14	100.0	456	6	CA995466
C 8	14	100.0	458	6	CA996608
C 9	14	100.0	459	5	BQ125427
C 10	14	100.0	460	5	BQ626540
C 11	14	100.0	461	5	BQ626708
C 12	14	100.0	531	7	CF266988
C 13	14	100.0	562	9	TA3A03P
C 14	14	100.0	570	8	AQ435225
C 15	14	100.0	572	7	CV193301
C 16	14	100.0	595	1	AJ716704
C 17	14	100.0	654	2	BF219779
C 18	14	100.0	667	8	BH586036
C 19	14	100.0	702	5	EX921917
C 20	14	100.0	720	2	BF485581
C 21	14	100.0	737	4	BJ591199
C 22	14	100.0	763	4	BJ600497
C 23	14	100.0	765	8	BH438319
C 24	14	100.0	766	4	BJ611184

C 25	14	100.0	775	4	BJ589511
C 26	14	100.0	783	8	BH473894
C 27	14	100.0	786	9	CG811919
C 28	14	100.0	811	7	CF468731
C 29	14	100.0	815	9	BX969025
C 30	14	100.0	817	9	CR209043
C 31	14	100.0	826	7	CK462079
C 32	14	100.0	830	9	CR053794
C 33	14	100.0	831	7	CK465310
C 34	14	100.0	832	9	EX997284
C 35	14	100.0	865	7	CK461721
C 36	14	100.0	945	4	BG121851
C 37	14	100.0	1305	8	CC206609
C 38	13.6	97.1	1950	5	CNS02EL2
C 39	13.2	94.3	1098	5	EX395152
C 40	13	92.9	117	8	BZ647319
C 41	13	92.9	130	7	W11332
C 42	13	92.9	267	8	AQ105334
C 43	13	92.9	292	1	AA061063
C 44	13	92.9	302	5	BW033379
C 45	13	92.9	307	9	CC611001

ALIGNMENTS

RESULT 1
AV269567/c
LOCUS
DEFINITION AV269567 RIKEN full-length enriched, adult male testis (DH10B) Mus
musculus CDNA clone 4930543P08 3', mRNA sequence.
ACCSSION AV269567
VERSION AV269567.1 GI:6257604
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 204)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,
Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,
Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tomimaga, N.,
Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A.,
Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gs.c.riken.jp/
Sasaki, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
Hayashizaki, Y.

TITLE
JOURNAL
COMMENT

Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Ktsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for


```

MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9585
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 Reverse.
FEATURES
Source
Location/Qualifiers
1. .324
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-a1-c-01-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATTACAGA; lens, CGATTACGA; eye anterior segment,
AATGCCGAT; optic nerve, CCATTAAAGTG; retina, CGCGG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."
```

ORIGIN

Query Match 100.0%; Score 14; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||

Db 295 GGCATCGTCAGTTG 282
|||||

RESULT 5
CB486369
LOCUS CB486369 438 bp mRNA linear EST 01-APR-2003
DEFINITION Omykrch001067 Oncorhynchus mykiss reproductive Oncorhynchus mykiss
cDNA, mRNA sequence.
ACCESSION CB486369.1 GI:29297595
VERSION CB486369.1
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 438)
AUTHORS GRASP Consortium, Davidson, W.S., Koop, B.F. and
http://web.uvic.ca/cbr/grasp
TITLE A survey of Salmo salar transcripts from high complexity cDNA

libraries
Unpublished (2002)
Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca
Centre for Biomedical Research, University of Victoria cDNA
preparation and sequencing: Roberto Alberto, Marianne
Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.
bioinformatics: Gordon D Brown.
Location/Qualifiers
1. .438
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/strain="Karl Vierke"
/db_xref="taxon:8022"
/clone_lib="Oncorhynchus mykiss reproductive"
/notes="vector: pcrt4topo; Library Creator: Kristian R von
Schalburg; Rainbow trout tissue contributors: Mountain
Trout Sales (Sooke, B.C.)"

ORIGIN

Query Match 100.0%; Score 14; DB 6; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||

Db 291 GGCATCGTCAGTTG 304
|||||

RESULT 6
CA995113
LOCUS CA995113 456 bp mRNA linear EST 07-JAN-2003
DEFINITION rg24c07.y1 Meloidogyne hapla J2 SL1 TOPO v1 Meloidogyne hapla cDNA
5', mRNA sequence.
ACCESSION CA995113
VERSION CA995113.1 GI:27539984
KEYWORDS EST.
SOURCE Meloidogyne hapla
ORGANISM Meloidogyne hapla
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE 1 (bases 1 to 456)
AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
Teagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter
at Washington University, St. Louis. J2 were provided by Dr.
Valerie Williamson of the University of California at Davis
(vmwilliamson@ucdavis.edu).
Putative full length read
The vector to vector length is 457
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .456
/organism="Meloidogyne hapla"

```

/mol_type="mRNA"
/db_xref="taxon:6305"
/dev_stage="J2"
/lab_host="DH10B"
/clone_lib="Meloiodogyne hapla J2 SL1 TOPO v1"
/notes=vector; pCRII-TOPO (Invitrogen); Site 1: EcoRI;
Site 2: EcoRI; The library was constructed by Claire
Murphy and Dr. James McCarter at Washington University,
St. Louis. Oligo(dT)-SL1 PCR based library. Meloiodogyne
hapla J2 cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into pCRII-TOPO(Invitrogen)
following the Topo TA cloning protocol. J2 were provided
by Dr. Valerie Williamson of University of California at
Davis (vmwilliamson@ucdavis.edu)."

ORIGIN
Query Match 100.0%; Score 14; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 286 GGCATCGTCAGTTG 299

RESULT 7
CA995466
LOCUS
DEFINITION
rg29f01.v1 Meloiodogyne hapla J2 SL1 TOPO v1 Meloiodogyne hapla cDNA
5', mRNA sequence.
ACCESSION
CA995466.1 GI:27540337
KEYWORDS
EST.
SOURCE
Meloiodogyne hapla
ORGANISM
Meloiodogyne hapla
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloiodogyninae; Meloiodogyne.
REFERENCE
1 (bases 1 to 456)
AUTHORS
McCarte,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter
at Washington University, St. Louis. J2 were provided by Dr.
Valerie Williamson of the University of California at Davis
(vmwilliamson@ucdavis.edu).
Putative full length read
The vector to vector length is 457
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..456
/organism="Meloiodogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/dev_stage="J2"
/lab_host="DH10B"
/clone_lib="Meloiodogyne hapla J2 SL1 TOPO v1"
/notes=vector; pCRII-TOPO (Invitrogen); Site 1: EcoRI;
Site 2: EcoRI; The library was constructed by Claire
Murphy and Dr. James McCarter at Washington University,
St. Louis. Oligo(dT)-SL1 PCR based library. Meloiodogyne
hapla J2 cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into pCRII-TOPO(Invitrogen)
following the Topo TA cloning protocol. J2 were provided
by Dr. Valerie Williamson of University of California at
Davis (vmwilliamson@ucdavis.edu)."

ORIGIN
Query Match 100.0%; Score 14; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
|||||
Db 286 GGCATCGTCAGTTG 299

RESULT 8
CA996608
LOCUS
DEFINITION
rg43e05.y1 Meloiodogyne hapla J2 SL1 TOPO v1 Meloiodogyne hapla cDNA
5', mRNA sequence.
ACCESSION
CA996608.1 GI:27541479
KEYWORDS
EST.
SOURCE
Meloiodogyne hapla
ORGANISM
Meloiodogyne hapla
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloiodogyninae; Meloiodogyne.
REFERENCE
1 (bases 1 to 458)
AUTHORS
McCarte,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter
at Washington University, St. Louis. J2 were provided by Dr.
Valerie Williamson of the University of California at Davis
(vmwilliamson@ucdavis.edu).
Putative full length read
The vector to vector length is 459
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..458
/organism="Meloiodogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/dev_stage="J2"
/lab_host="DH10B"
/clone_lib="Meloiodogyne hapla J2 SL1 TOPO v1"
/notes=vector; pCRII-TOPO (Invitrogen); Site 1: EcoRI;
Site 2: EcoRI; The library was constructed by Claire
Murphy and Dr. James McCarter at Washington University,
St. Louis. Oligo(dT)-SL1 PCR based library. Meloiodogyne
hapla J2 cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into pCRII-TOPO(Invitrogen)
following the Topo TA cloning protocol. J2 were provided
by Dr. Valerie Williamson of University of California at
Davis (vmwilliamson@ucdavis.edu)."

FEATURES
source
1..458
/organism="Meloiodogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/dev_stage="J2"
/lab_host="DH10B"
/clone_lib="Meloiodogyne hapla J2 SL1 TOPO v1"
/notes=vector; pCRII-TOPO (Invitrogen); Site 1: EcoRI;
Site 2: EcoRI; The library was constructed by Claire
Murphy and Dr. James McCarter at Washington University,
St. Louis. Oligo(dT)-SL1 PCR based library. Meloiodogyne
hapla J2 cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into pCRII-TOPO(Invitrogen)
following the Topo TA cloning protocol. J2 were provided
by Dr. Valerie Williamson of University of California at
Davis (vmwilliamson@ucdavis.edu)."

```

Query Match 100.0%; Score 14; DB 6; Length 458;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14

Db 288 GGCATCGTCAGTTG 301

RESULT 9
 BQ125427
 LOCUS
 DEFINITION rc68b06.v1 Meloidogyne hapla egg SL1 TOPO v1 Meloidogyne hapla
 cDNA mRNA sequence.
 ACCESSION BQ125427
 VERSION BQ125427.1 GI:20199338
 KEYWORDS EST.
 SOURCE Meloidogyne hapla
 ORGANISM Meloidogyne hapla
 Eukaryote; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
 REFERENCE 1 (bases 1 to 459)
 AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
 Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
 Tagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 CONTACT: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 The library was constructed by Claire Murphy and Dr. James McCarter
 at Washington University, St. Louis. Eggs were provided by Dr.
 David Bird of North Carolina State University, Raleigh, NC
 (david_bird@ncsu.edu). DNA Sequencing by: Washington University
 Genome Sequencing Center

Putative full length read
 The vector to vector length is 460
 Seq primer: SL1 primer.

FEATURES
 source
 Location/Qualifiers
 1..459
 /organism="Meloidogyne hapla"
 /mol_type="mRNA"
 /db_xref="taxon:6305"
 /dev_stage="eggs"
 /lab_host="DH10b"
 /clone_lib="Meloidogyne hapla egg SL1 TOPO v1"
 /note="Vector: pCRII-TOPO (Invitrogen); Site1: EcoRI;
 Site2: EcoRI; The library was constructed by Claire
 Murphy and Dr. James McCarter at Washington University,
 St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne
 hapla egg cDNA PCR products of size >400 nucleotides
 containing SL1 on the 5' end and oligo(dT) on the 3' end
 were non-directionally cloned into pCRII-TOPO(Invitrogen)
 following the Topo TA cloning protocol. Eggs were provided
 by Dr. David Bird of North Carolina State University,
 Raleigh, NC (david_bird@ncsu.edu)"

ORIGIN

Query Match 100.0%; Score 14; DB 5; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14

Db 289 GGCATCGTCAGTTG 302

RESULT 10
 BQ626540
 LOCUS
 DEFINITION pz1le03.y1 Pratylenchus penetrans mixed stage SL1 TOPO v1
 Pratylenchus penetrans cDNA 5', mRNA sequence.
 ACCESSION BQ626540
 VERSION BQ626540.1 GI:21653718
 KEYWORDS EST.
 SOURCE Pratylenchus penetrans
 ORGANISM Pratylenchus penetrans

Eukaryote; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Pratylenchidae; Pratylenchinae; Pratylenchus.
 REFERENCE 1 (bases 1 to 460)
 AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
 Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
 Tagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 CONTACT: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 The library was constructed by Claire Murphy and Dr. James McCarter
 at Washington University, St. Louis. RNA was provided by Andrew
 Kloeck of Divergence Inc., St. Louis, MO.

Putative full length read
 The vector to vector length is 461
 Seq primer: SL1 primer.

FEATURES
 source
 Location/Qualifiers
 1..460
 /organism="Pratylenchus penetrans"
 /mol_type="mRNA"
 /db_xref="taxon:45929"
 /dev_stage="mixed stage"
 /lab_host="DH10B (Invitrogen)"
 /clone_lib="Pratylenchus penetrans mixed stage SL1 TOPO
 v1"
 /note="Vector: pCRII-TOPO (Invitrogen); Site1: EcoRI;
 Site2: EcoRI; The library was constructed by Claire
 Murphy and Dr. James McCarter at Washington University,
 St. Louis. Oligo(dT)-SL1 PCR based library. cDNA PCR
 products of size >400 nucleotides containing SL1 on the 5'
 end and oligo(dT) on the 3' end were non-directionally
 cloned into pCRII-TOPO(Invitrogen) following the Topo TA
 cloning protocol. RNA was provided by Andrew Kloeck of
 Divergence, Inc., St. Louis, MO."

ORIGIN

Query Match 100.0%; Score 14; DB 5; Length 460;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14

Db 290 GGCATCGTCAGTTG 303

RESULT 11
 BQ626708
 LOCUS
 DEFINITION pz13h07.y1 Pratylenchus penetrans mixed stage SL1 TOPO v1
 Pratylenchus penetrans cDNA 5', mRNA sequence.

```

ACCESSION   BQ626708
VERSION     BQ626708.1  GI:21653886
SOURCE      EST.
ORGANISM    Pratylenchus penetrans
            Pratylenchus Penetrans
            Tylenchoidea; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Pratylenchidae; Pratylenchinae; Pratylenchus.

REFERENCE   1 (bases 1 to 461)
AUTHORS    McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Stepcoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
            The Washington Univ. Nematode EST Project, 1999
            Unpublished (1999)
            Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. RNA was provided by Andrew
            Kloek of Divergence Inc., St. Louis, MO.
            Putative full length read
            The vector to vector length is 462
            Seq primer: SL1 primer.
            Location/Qualifiers
            1..461
               /organism="Pratylenchus penetrans"
               /mol_type="mRNA"
               /db_xref="taxon:45929"
               /dev_stage="mixed stage"
               /lab_host="DH10B (Invitrogen)"
               /clone_lib="Pratylenchus penetrans mixed stage SL1 TOPO
               v1"
               /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
               Site_2: EcoRI; The library was constructed by Claire
               Murphy and Dr. James McCarter at Washington University,
               St. Louis. Oligo(dT)-SL1 PCR based library. cDNA PCR
               products of size >400 nucleotides containing SL1 on the 5'
               end and oligo(dT) on the 3' end were non-directionally
               cloned into pCRII-TOPO(Invitrogen) following the Topo TA
               cloning protocol. RNA was provided by Andrew Kloek of
               Divergence, Inc., St. Louis, MO."

FEATURES             source
     source
     1..461
        Location/Qualifiers
        1..531
           /organism="Toxoplasma gondii"
           /mol_type="mRNA"
           /strain="tachyzoite"
           /db_xref="taxon:5811"
           /clone="TgESTzj21a08.y1"
           /dev_stage="Tachyzoite"
           /lab_host="Electroten Blue cells (Stratagene)"
           /clone_lib="TgMAS tachyzoite cDNA Library"
           /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
           XhoI; The cDNA library was constructed by Keliang Tang,
           and Robert Cole at Washington University. cDNA was
           synthesized from Poly(A)+ mRNA using an oligo-d(T) primer
           containing a XhoI site. Following second strand synthesis,
           EcoRI adapters were ligated to the cDNA, and products were
           size-selected on sephacryl S500. The cDNA were
           directionally cloned into the EcoRI/XhoI prepared
           pBluescript II SK+ vector, and electroporated into
           Electroten Blue cells (Stratagene). The library may
           contain a small percentage of host or bacterial
           contaminants."

ORIGIN
Query Match      100.0%; Score 14; DB 7; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 40 GGCATCGTCAGTTG 27

RESULT 13
TA43A03P/c
LOCUS
DEFINITION     T. brucei sheared genomic DNA clone 43a03, forward sequence,
                genomic survey sequence.
ACCESSION      AL454668
VERSION        AL454668.1  GI:11856292
KEYWORDS       GSS.
SOURCE         Trypanosoma brucei
               Trypanosoma brucei
               Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
               Trypanosoma.
REFERENCE      1 (bases 1 to 562)
AUTHORS       Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
               Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
               Melville,S.E., Rajadream,M.A. and Barrell,B.G.
               Direct Submission
               Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
               project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
               Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
               nh@sanger.ac.uk
               Constructed at the Institute for Genomic Research (TIGR),
               Rockville, MD. Genomic DNA isolated from a cloned population of

TITLE          JOURNAL
COMMENT

```

Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org
 Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source
 1. .562
 /organism="Trypanosoma brucei"
 /mol_type="genomic DNA"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clones="43a03"

ORIGIN

Query Match 100.0%; Score 14; DB 9; Length 562;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
 |||||
 Db 148 GGCATCGTCAGTTG 135

RESULT 14

LOCUS A0435225/c
 DEFINITION HS_5140_B2_E07_SP6E RPCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=716 Col=14 Row=J, genomic survey sequence.
 A0435225

ACCESSION A0435225.1 GI:4546564

VERSION GSS.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 570)
 Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

9380589

1049764

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallaceu.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 716 row: J column: 14

Seq primer: SP6

Class: BAC ends

High quality sequence stop: 570.

FEATURES

source
 1. .570
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clones="Plate=716 Col=14 Row=J"
 /sex="male"

/clone_lib="RPCI-11 Human Male BAC Library"

/notes="vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

ORIGIN

Query Match 100.0%; Score 14; DB 8; Length 570;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
 |||||
 Db 342 GGCATCGTCAGTTG 329

RESULT 15

CV193301/c

LOCUS

DEFINITION CV193301 572 bp mRNA linear EST 15-SEP-2004
 SNEStab28h01.y1 Sarcocystis neuropa cDNA 5', mRNA sequence.
 Library Sarcocystis neuropa cDNA 5', mRNA sequence.

ACCESSION CV193301

VERSION CV193301.1 GI:52122138

KEYWORDS EST.

SOURCE Sarcocystis neuropa

ORGANISM Sarcocystis neuropa

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Sarcocystis.

1 (bases 1 to 572)

REFERENCE

AUTHORS

Howe D.K., Stauffer S., Tang K., Sibley L.D., Clifton S., Marra M., Hillier L., Pape D., Martin J., Wylie T., Theising B., Bowers Y., Gibbons M., Ritter E., McCann R., Blistain A., Bennet J., Schmitt A., Ronko I., Tsagarisshvili R., Fedele M., Belaygorod L., Franklin C., Carr L.M., Grow A., Maguire L., Wadkins J., Richey J., Waterston R. and Wilson R.

Sarcocystis neuropa EST project

Unpublished (2000)

CONTACT: Daniel K. Howe

Sarcocystis neuropa EST project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Contact Daniel K. Howe (dkhowe2@pop.uky.edu) for further

information relating to organism, libraries, or clone availability.

Sequenced by Washington University Genome Sequencing Center

Seq primer: -21UPPOT

High quality sequence stop: 572.

Location/Qualifiers

1. .572

/organism="Sarcocystis neuropa"

/mol_type="mRNA"

/strain="SN4"

/db_xref="taxon:42890"

/lab_stage="merozoite"

/dev_host="GC10"

/clone_lib="Sarcocystis neuropa merozoite UK CSN4 1 CDNA library"

/notes="Vector: pDNR-LIB; Site_1: GGCGCTCGGCC; Site_2: GGCGCTCGGCC; Library constructed by: Daniel K. Howe and Michelle R. Yeagan Total RNA was isolated from culture-derived merozoites of Sarcocystis neuropa strain SN4. cDNA was synthesized using the template-switching and long-distance PCR method (SMART cDNA library construction kit, BD Biosciences). The amplified cDNA fragments were digested with SfiI, size fractionated, and ligated into SfiI-digested pDNR-LIB vector."

ORIGIN

Query Match 100.0%; Score 14; DB 7; Length 572;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
 |||||
 Db 124 GGCATCGTCAGTTG 111

RESULT 16
 AJ716704
 LOCUS
 DEFINITION AJ716704 Triticum turgidum subsp. durum etiolated seedling 20 days
 Triticum turgidum subsp. durum cDNA clone 05257R, mRNA sequence.
 AJ716704
 ACCESSION AJ716704.1 GI:49603286
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Triticum turgidum subsp. durum (durum wheat)
 Triticum turgidum subsp. durum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Triticum.
 1 (bases 1 to 595)
 Cifarelli,R.A., D'Onofrio,O. and Lauria,G.
 Expressed Sequences Tags (ESTs) library from totipotent cDNA of
 durum wheat
 Unpublished (2003)
 Contact: Cifarelli RA
 Biotechnology
 Metapontum Agrobios
 S.S. Jonica 106 km 448.2, 75010 Metaponto (MT), Italy.
 Location/Qualifiers
 1. 595
 /organism="Triticum turgidum subsp. durum"
 /mol_type="mRNA"
 /cultivar="Ofanto"
 /sub_species="durum"
 /db_xref="taxon:4567"
 /clone="05257R"
 /tissue_type="etiolated seedling"
 /dev_stage="20 days"
 /clone_lib="Triticum turgidum subsp. durum etiolated
 seedling 20 days"

ORIGIN
 Query Match 100.0%; Score 14; DB 1; Length 595;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
 |||||
 Db 414 GGCATCGTCAGTTG 427

RESULT 17
 BF219779/c
 LOCUS
 DEFINITION BF219779 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822421 5',
 mRNA sequence.
 ACCESSION BF219779
 VERSION BF219779.1 GI:11125873
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 654)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: IRALI row: i column: 04
 High quality sequence stop: 613.
 Location/Qualifiers
 1. 654

FEATURES

source
 1. 654
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2822421"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 7"
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 100.0%; Score 14; DB 2; Length 654;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
 |||||
 Db 619 GGCATCGTCAGTTG 606

RESULT 18

BH586036/c
 LOCUS
 DEFINITION BH586036 BOGR Brassica oleracea genomic clone BOGRS44, genomic
 survey sequence.
 ACCESSION BH586036
 VERSION BH586036.1 GI:117838494
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 667)
 Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 Other GSs: BOGRS44TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1. 667
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone_lib="BOGR"
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOS1 using BstXI linkers"

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

source
 1. 667
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone_lib="BOGR"
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match 100.0%; Score 14; DB 8; Length 667;


```

Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 542 GGCATCGTCAGTTG 529

RESULT 19
BX921917/c
LOCUS
DEFINITION BX921917 Sus Scrofa library (scan) linear EST 07-MAY-2004
scan0012d.j.07 5prim, mRNA sequence.
ACCESSION BX921917
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 702)
Bonnef,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villegier,S.,
Soares,M., Bonaldo,F. and Hatey,F.
A Pig Normalised Multi-Tissue cDNA Library
Unpublished (2003)
COMMENT
Contact: Tosser-Klopp G
Genetique Animale
Institut National de la Recherche Agronomique
Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
cedex, FRANCE
Tel: 33 (0) 5.61.28.51.14
Fax: 33 (0) 5.61.28.53.08
Email: tosser@toulouse.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0012 row: j column: 7.

FEATURES
source
Location/Qualifiers
1..702
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="scan0012d.j.07"
/tissue_type="mixed"
/dev_stage="from embryos to adults"
/clone_lib="Sus Scrofa library (scan)"
/notes="tissues: adipose tissue, brain, kidney, liver,
muscle, ovary, testis, heart, hypothalamus, pancreas,
skin, spleen, thymus, placenta, pituitary gland, seminal
vesicle, small intestine, uterus, adrenals, bulbo urethral
gland, cerebral trunk, epididymis, female gonad,
gall-bladder, hippocampus, large intestine, male gonad,
melanocytes, stomach, udder"

ORIGIN
Query Match 100.0%; Score 14; DB 5; Length 702;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 166 GGCATCGTCAGTTG 153

RESULT 20
BF485581/c
LOCUS
DEFINITION BF485581 AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT19621 5, mRNA sequence.
ACCESSION BF485581
VERSION
KEYWORDS
SOURCE
Drosophila melanogaster (fruit fly)

Query Match 100.0%; Score 14; DB 5; Length 702;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 166 GGCATCGTCAGTTG 153

RESULT 20
BF485581/c
LOCUS
DEFINITION BF485581 AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT19621 5, mRNA sequence.
ACCESSION BF485581
VERSION
KEYWORDS
SOURCE
Drosophila melanogaster (fruit fly)

Query Match 100.0%; Score 14; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 357 GGCATCGTCAGTTG 344

RESULT 21
BJ591199/c
LOCUS
DEFINITION BJ591199 normalized full length cDNA library, chloronemata,
caulonemata and malformed buds Physcomitrella patens subsp. patens
cDNA clone pphb43b12 3', mRNA sequence.
ACCESSION BJ591199
VERSION BJ591199.1 GI:37833187
KEYWORDS
SOURCE
ORGANISM
Physcomitrella patens subsp. patens
Physcomitrella patens subsp. patens
Eukaryota; Viridiplantae; Streptophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 737)
Nishiyama,T., Fujita,T., Shin-i,T., Seki,M., Nishide,H.,
Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
Kohara,Y. and Hasebe,M.
Comparative genomics of Physcomitrella patens gametophytic
transcriptome and Arabidopsis thaliana: implication for land plant
evolution
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 720)
Stapleton,M., Brokstein,P., Hong,L., Agbavani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celniker,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
COMMENT
On Dec 6, 2000 this sequence version replaced gi:11568918.
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd. Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003593: arm:3L [20799555,21112669]
estimated-cyto:78A2-78C7: 04/09/2001
Plate: AT.196 row: B column: 9
High quality sequence stop: 647.

FEATURES
source
Location/Qualifiers
1..720
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT19621"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/notes="Organ: ADULT testes; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

ORIGIN
Query Match 100.0%; Score 14; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCATCGTCAGTTG 14
    |||||
Db 357 GGCATCGTCAGTTG 344

RESULT 21
BJ591199/c
LOCUS
DEFINITION BJ591199 normalized full length cDNA library, chloronemata,
caulonemata and malformed buds Physcomitrella patens subsp. patens
cDNA clone pphb43b12 3', mRNA sequence.
ACCESSION BJ591199
VERSION BJ591199.1 GI:37833187
KEYWORDS
SOURCE
ORGANISM
Physcomitrella patens subsp. patens
Physcomitrella patens subsp. patens
Eukaryota; Viridiplantae; Streptophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 737)
Nishiyama,T., Fujita,T., Shin-i,T., Seki,M., Nishide,H.,
Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
Kohara,Y. and Hasebe,M.
Comparative genomics of Physcomitrella patens gametophytic
transcriptome and Arabidopsis thaliana: implication for land plant
evolution
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

```

MEDLINE PUBMED COMMENT	22709184 12808149 Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- GAGAGAGAGGATCCACCCCTGGAGAGTGTTCCTTCGATCGATGCTCGAGACGATGCTCGAGACCGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCCGAGCTCGAATTCGTCGAGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13- 14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database of Physcomitrella EST clones is available at the PHYSCOBASE (http://moss.nibb.ac.jp).
FEATURES source	1. .737 Location/Qualifiers /organism="Physcomitrella patens subsp. patens" /mol_type="mRNA" /sub_species="patens" /db_xref="taxon:145481" /clone="pphb3b12" /tissue_type="mixture of chloronemata, caulonemata and malformed buds" /clone_lib="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"
ORIGIN	Query Match 100.0%; Score 14; DB 4; Length 737; Best Local Similarity 100.0%; Pred. No. 1.5e+03; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GGATCGTCAGTTG 14
Db	690 GGATCGTCAGTTG 677
RESULT 22 BJ600497/c LOCUS	763 bp mRNA linear EST 22-OCT-2003
DEFINITION	BJ600497 normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata Physcomitrella patens subsp. patens cDNA clone pphn29e17 3', mRNA sequence.
ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BJ600497 BJ600497.1 GI:37842489 EST. Physcomitrella patens subsp. patens Physcomitrella patens subsp. patens Eukaryota; Viridiplantae; Streptophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 763) Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M. Comparative genomics of Physcomitrella patens gametophytic evolution transcriptome and Arabidopsis thaliana: implication for land plant evolution
REFERENCE AUTHORS TITLE	Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003) 22709184 PUBMED
JOURNAL MEDLINE PUBMED	22709184 12808149
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- GAGAGAGAGGATCCACCCCTGGAGAGTGTTCCTTCGATCGATGCTCGAGACGATGCTCGAGACCGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCCGAGCTCGAATTCGTCGAGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13- 14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database of Physcomitrella EST clones is available at the PHYSCOBASE (http://moss.nibb.ac.jp).
FEATURES source	1. .763 Location/Qualifiers /organism="Physcomitrella patens subsp. patens" /mol_type="mRNA" /sub_species="patens" /db_xref="taxon:145481" /clone="pphn29e17" /tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata" /clone_lib="normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata"
ORIGIN	Query Match 100.0%; Score 14; DB 4; Length 763; Best Local Similarity 100.0%; Pred. NO. 1.5e+03; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GGATCGTCAGTTG 14
Db	718 GGATCGTCAGTTG 705
RESULT 23 BH438319 LOCUS	765 bp DNA linear GSS 12-DEC-2001
DEFINITION	BOGH102TR BOGH Brassica oleracea genomic clone BOGH102, genomic survey sequence.
ACCESSION VERSION KEYWORDS SOURCE ORGANISM	BH438319 BH438319 BH438319.1 GI:17624040 GSS. Brassica oleracea Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 765) Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M. Whole genome shotgun sequencing of Brassica oleracea Unpublished (2001) Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TR

```

FEATURES
  source
    Class: sheared ends
    Location/Qualifiers
      1..765
        /organism="Brassica oleracea"
        /mol_type="genomic DNA"
        /strain="TO1000DH3"
        /db_xref="taxon:3712"
        /clone="BOGH102"
        /clone_lib="BOGH"
        /notes="vector: pSHO1; Site 1: BstXI; 2-3 kb sheared
        genomic DNA inserted into pSHO1 using BstXI linkers"

ORIGIN
  Query Match      100.0%; Score 14; DB 8; Length 765;
  Best Local Similarity 100.0%; Pred. NO. 1.5e+03;
  Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 GGCATCGTCAGTTG 14
      |||||
  Db 315 GGCATCGTCAGTTG 328

RESULT 24
BJ611184/c
LOCUS
DEFINITION
  BJ611184 766 bp mRNA linear EST 22-OCT-2003
  BJ611184 normalized full length cDNA library, chloronemata,
  caulonemata and rhizoid-like protonemata Physcomitrella patens
  subsp. patens cDNA clone pphn14e10 3', mRNA sequence.
  BJ611184
  VERSION
  SOURCE
  ORGANISM
    Physcomitrella patens subsp. patens
    Physcomitrella patens subsp. patens
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
    Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
  REFERENCE
  1 (bases 1 to 766)
    Nishiyama,T., Fujita,T., Shin-i.T., Seki,M., Nishide,H.,
    Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
    Kohara,Y. and Hasebe,M.
  TITLE
    Comparative genomics of Physcomitrella patens gametophytic
    transcriptome and Arabidopsis thaliana: implication for land plant
    evolution
  JOURNAL
  MEDLINE
  PUBMED
  COMMENT
    Contact: Tadasu Shin-i
    Center For Genetic Resource Information
    National Institute of Genetics
    1111 Yata, Mishima, Shizuoka 411-8540, Japan
    Tel: 81-559-81-6856
    Fax: 81-559-81-6855
    Email: tshini@genes.nig.ac.jp
    A backbone of the vector is basically from pBluescript II (KS),
    that was in vivo excised from a 1-PLC phage vector (Carninci et al.
    2001). 5' end of the cDNA that was digested with XhoI was ligated
    to SalI site of the vector and the 3' end including polyA tail was
    ligated to BamHI site of the
    vector(5'- GAGAGAGAGAGATCCACCCCTGGAGAGATTTTTTTTTTTT-3' was
    used as a 1st 3' primer, and
    5'-GGTTCTCGATCGTCGTTCTCCAGACGATGACTCGAGAACCGNNNN-3' as 2nd
    5'-hairpin primer giving the following 5' boarder sequence,
    AGGCAATCGCCGAGCTCGAATTCGAGAACCG). cDNA insert could be
    amplified with conventional T7 and T3 primers. This full-length
    cDNA library was generated according to the method described in
    Nishiyama et al. (2003).
    The BCDATG medium for 13- 14 days under the continuous light.
    These clones are available from RIKEN Bio Resource Center
    (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
    of Physcomitrella EST clones is available at the PHYSCObase
    (http://moss.nibb.ac.jp).
    Location/Qualifiers
      1..766
        /organism="Brassica oleracea"
        /mol_type="genomic DNA"

```

```

/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphn14e10"
/tissue_type="mixture of chloronemata, caulonemata and
rhizoid-like protonemata"
/rhizoid-lib="normalized full length cDNA library,
chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN
  Query Match      100.0%; Score 14; DB 4; Length 766;
  Best Local Similarity 100.0%; Pred. NO. 1.5e+03;
  Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 GGCATCGTCAGTTG 14
      |||||
  Db 718 GGCATCGTCAGTTG 705

RESULT 25
BJ589511/c
LOCUS
DEFINITION
  BJ589511 775 bp mRNA linear EST 22-OCT-2003
  BJ589511 normalized full length cDNA library, chloronemata,
  caulonemata and malformed buds Physcomitrella patens subsp. patens
  cDNA clone pphb39112 3', mRNA sequence.
  BJ589511
  VERSION
  SOURCE
  ORGANISM
    Physcomitrella patens subsp. patens
    Physcomitrella patens subsp. patens
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
    Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
  REFERENCE
  1 (bases 1 to 775)
    Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
    Kohara,Y. and Hasebe,M.
  TITLE
    Comparative genomics of Physcomitrella patens gametophytic
    transcriptome and Arabidopsis thaliana: implication for land plant
    evolution
  JOURNAL
  MEDLINE
  PUBMED
  COMMENT
    Contact: Tadasu Shin-i
    Center For Genetic Resource Information
    National Institute of Genetics
    1111 Yata, Mishima, Shizuoka 411-8540, Japan
    Tel: 81-559-81-6856
    Fax: 81-559-81-6855
    Email: tshini@genes.nig.ac.jp
    A backbone of the vector is basically from pBluescript II (KS),
    that was in vivo excised from a 1-PLC phage vector (Carninci et al.
    2001). 5' end of the cDNA that was digested with XhoI was ligated
    to SalI site of the vector and the 3' end including polyA tail was
    ligated to BamHI site of the
    vector(5'- GAGAGAGAGAGATCCACCCCTGGAGAGATTTTTTTTTTTT-3' was
    used as a 1st 3' primer, and
    5'-GGTTCTCGATCGTCGTTCTCCAGACGATGACTCGAGAACCGNNNN-3' as 2nd
    5'-hairpin primer giving the following 5' boarder sequence,
    AGGCAATCGCCGAGCTCGAATTCGAGAACCG). cDNA insert could be
    amplified with conventional T7 and T3 primers. This full-length
    cDNA library was generated according to the method described in
    Nishiyama et al. (2003).
    The BCDATG medium for 13- 14 days under the continuous light.
    These clones are available from RIKEN Bio Resource Center
    (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
    of Physcomitrella EST clones is available at the PHYSCObase
    (http://moss.nibb.ac.jp).
    Location/Qualifiers
      1..775
        /organism="Physcomitrella patens subsp. patens"
        /mol_type="mRNA"

```

```
/sub_species="patens"  
/db_xref="taxon:145481"  
/clone="pphb39112"  
/tissue_type="mixture of chloronemata, caulonemata and  
malformed buds"  
/clone_lib="normalized full length cDNA library,  
chloronemata, caulonemata and malformed buds"
```

ORIGIN

```
Query Match      100.0%; Score 14; DB 4; Length 775;  
Best Local Similarity 100.0%; Pred. NO. 1.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 GGCATCGTCAGTTG 14  
        |||||  
Db      691 GGCATCGTCAGTTG 678
```

Search completed: March 11, 2005, 13:01:15
Job time : 188.257 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 23.4069 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-19

Perfect score: 16

Sequence: 1 cggcatcgctcagttgc 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002as:*

7: Geneseq2002bs:*

8: Geneseq2003as:*

9: Geneseq2003bs:*

10: Geneseq2003cs:*

11: Geneseq2003ds:*

12: Geneseq2004as:*

13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	100.0	16	3	Aaz36119 Primer de
2	16	100.0	18	3	Aaz36120 Primer de
3	16	100.0	31	3	Aaz36112 Primer de
4	16	100.0	31	3	Aaz36113 Primer de
5	16	100.0	1489	3	Aaz36101 Nucleic a
6	15	93.8	319	10	Adk56969 Plant DNA
7	15	93.8	319	10	Adk55125 Plant DNA
8	15	93.8	340	10	Adk53158 Plant DNA
9	15	93.8	368	10	Adk56429 Plant DNA
10	14.4	90.0	60	6	Abn40186 Human spl
11	14.4	90.0	387	9	Ach42666 Human foe
12	14.4	90.0	498	9	Ach32955 Human end
13	14.4	90.0	664	10	Adb68745 Pseudomon
14	14.4	90.0	930	8	ACA40338 Prokaryot
15	14.4	90.0	936	8	ACA38540 Prokaryot
16	14.4	90.0	1056	13	Adk46729 Bacterial
17	14.4	90.0	1137	6	Abz13613 Arabidops
18	14.4	90.0	1301	3	Aac33636 Arabidops
19	14.4	90.0	1362	13	Adt46111 Bacterial
20	14.4	90.0	1392	8	ACA36906 Prokaryot

C 21	14.4	90.0	1563	10	ACF70905	Acf70905 Photorhab
C 22	14.4	90.0	1632	2	AAQ96251	Aaq96251 Fructosyl
C 23	14.4	90.0	1714	3	AAQ77952	Aaq77952 Human can
C 24	14.4	90.0	1734	13	ADS47843	Ads47843 Bacterial
C 25	14.4	90.0	1797	10	AAD57635	Aad57635 Rice dise
26	14.4	90.0	1821	4	AAF31627	Aaf31627 Mycobacte
27	14.4	90.0	1980	4	AAF31642	Aaf31642 Mycobacte
28	14.4	90.0	2220	4	AAF31643	Aaf31643 Mycobacte
29	14.4	90.0	2328	11	ABD07125	Abd07125 Pseudomon
C 30	14.4	90.0	2510	4	AAC84624	Aac84624 Human CUL
C 31	14.4	90.0	2511	11	ADI31609	Adi31609 Human cDN
C 32	14.4	90.0	2511	12	ADJ62794	Adj62794 Human cDN
C 33	14.4	90.0	4230	11	ABD07107	Abd07107 Pseudomon
34	14.4	90.0	4941	11	ABD07137	Abd07137 Pseudomon
35	14.4	90.0	19547	4	AAS59601	Aas59601 Propionib
36	14.4	90.0	19547	8	ACF64530	Acf64530 Propionib
37	14.4	90.0	22934	4	AAS59613	Aas59613 Propionib
38	14.4	90.0	22934	8	ACF64542	Acf64542 Propionib
C 39	14.4	90.0	25860	11	ADR20887	Adr20887 Photorhab
40	14.4	90.0	39052	4	ABL02664	Abi02664 Drosophil
41	14.4	90.0	110000	4	AAI99682_05	Continuation (6 of
42	14.4	90.0	110000	4	AAI99682_06	Continuation (7 of
43	14.4	90.0	110000	4	AAI99682_32	Continuation (33 o
44	14.4	90.0	110000	4	AAI99683_05	Continuation (6 of
45	14.4	90.0	110000	4	AAI99683_06	Continuation (7 of

ALIGNMENTS

RESULT 1

AAZ36119
ID AAZ36119 standard; DNA; 16 BP.

XX AC AAZ36119;

DT 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX WO9955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

XX Escherichia coli, particularly serotype O157:H7, used for detecting these

XX bacteria in food.

XX Claim 5; Page 27; 48pp; French.

XX AAZ36103-27 represent fragments derived from nucleic acid sequences

XX specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are

XX derived from two sequences. The first (AAZ36101) is 99.9% homologous to

XX the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and

XX 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).

XX The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 16 BP; 2 A; 5 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 1 CGGCATCGTCAGTTGC 16

RESULT 2
 AAZ36120
 ID AAZ36120 standard; DNA; 18 BP.

XX AAZ36120;

XX 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 XX PCR primer; probe; ss.

XX Synthetic.
 XX Escherichia coli.

XX WO9955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

XX AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 18 BP; 3 A; 5 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 3; Length 18;

Best Local Similarity 100.0%; Pred. No. 47;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 2 CGGCATCGTCAGTTGC 17

RESULT 3
 AAZ36112

ID AAZ36112 standard; DNA; 31 BP.

XX AAZ36112;

XX 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 XX PCR primer; probe; ss.

XX Synthetic.
 XX Escherichia coli.

XX WO9955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

XX AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 6 CGGCATCGTCAGTTGC 21

RESULT 4

AAZ36113

ID AAZ36113 standard; DNA; 31 BP.

```

XX AC AAZ36113;
XX DT 11-FEB-2000 (first entry)
XX DE Primer derived from a nucleic acid sequence specific to EHEC.
XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
XX KW PCR primer; probe; ss.
XX OS Synthetic.
XX OS Escherichia coli.
XX PN WO9955908-A2.
XX PD 04-NOV-1999.
XX PF 27-APR-1999; 99WO-FR001000.
XX PR 28-APR-1998; 98FR-00005329.
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX PI Frechon DTM, Laure FC, Thierry D;
XX DR WPI; 2000-013443/01.
XX DE New nucleic acid containing sequences specific to enterohemorrhagic
XX PT Escherichia coli, particularly serotype O157:H7, used for detecting these
XX PT bacteria in food.
XX PS Claim 5; Page 27; 48pp; French.
XX AC AAZ36103-27 represent fragments derived from nucleic acid sequences
XX CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
XX CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
XX CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
XX CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
XX CC The second sequence (AAZ36102) is associated with the presence of
XX CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
XX CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
XX CC for virulence proteins of Shigella flexneri. Both sequences are of
XX CC plasmid origin. The fragments are used as PCR primers and probes for the
XX CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
XX CC in human or animal samples, foods or the environment. The fragments are
XX CC also useful for epidemiological studies
XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 10 CGGCATCGTCAGTTGC 25

RESULT 5
AAZ36101
ID AAZ36101 standard; DNA; 1489 BP.
XX AC AAZ36101;
XX DT 11-FEB-2000 (first entry)
XX DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.
XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
XX KW IS91; ds.
XX OS Escherichia coli.

```

```

XX PN WO9955908-A2.
XX PD 04-NOV-1999.
XX PF 27-APR-1999; 99WO-FR001000.
XX PR 28-APR-1998; 98FR-00005329.
XX PA (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX PI Frechon DTM, Laure FC, Thierry D;
XX DR WPI; 2000-013443/01.
XX DE New nucleic acid containing sequences specific to enterohemorrhagic
XX PT Escherichia coli, particularly serotype O157:H7, used for detecting these
XX PT bacteria in food.
XX PS Claim 1; Fig 1; 48pp; French.
XX CC The present sequence is specific to enterohemorrhagic Escherichia coli
XX CC (EHEC). The sequence is 99.9% homologous to the katP gene of E. coli
XX CC O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
XX CC homologous with IS91 of E. coli (nucleotides 1-406 of the present
XX CC sequence). The present sequence is of plasmid origin. Fragments of the
XX CC present sequence are used, as probes and primers, for detection of E.
XX CC coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
XX CC animal samples, foods or the environment. The fragments are also useful
XX CC for epidemiological studies
XX SQ Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;

Query Match 100.0%; Score 16; DB 3; Length 1489;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 396 CGGCATCGTCAGTTGC 411

RESULT 6
ADK56969
ID ADK56969 standard; DNA; 319 BP.
XX AC ADK56969;
XX DT 06-MAY-2004 (first entry)
XX DE Plant DNA sequence which confers altered metabolic characteristic #4352.
XX KW altered metabolic characteristic; plant; acid metabolism;
XX KW alcohol metabolism; fatty acid metabolism;
XX KW branched fatty acid metabolism; alkaloid metabolism;
XX KW amino acid metabolism; ester metabolism; glyceride metabolism;
XX KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX OS Unidentified.
XX PN WO2003020936-A1.
XX PD 13-MAR-2003.
XX PF 30-AUG-2002; 2002WO-US027884.
XX PR 31-AUG-2001; 2001US-0316471P.
XX PA (DOWC ) DOW CHEM CO.
XX PA (DOWC ) DOW AGROSCIENCES LLC.

```

```

XX PT Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
PI PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX XX WPI; 2003-313091/30.
XX
XX PT Novel genes that confer altered metabolic characteristics in Nicotiana
PI PT benthiana plants, useful for altering the levels of metabolites e.g.
XX XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX PS Claim 1; SEQ ID NO 4352; 2576pp; English.
XX
XX CC The invention comprises DNA sequences which confer an altered metabolic
CC CC characteristic when they are expressed in a plant. The DNA sequences of
CC CC the invention are useful for producing plants with an altered metabolic
CC CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC CC altered glyceride metabolism, altered phenolic metabolism, altered
CC CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC CC invention may be used to provide disease resistance in a plant and gene
CC CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC CC DNA sequence of the invention.
XX
XX SQ Sequence 319 BP; 82 A; 77 C; 93 G; 67 T; 0 U; 0 Other;
XX
XX Query Match 93.8%; Score 15; DB 10; Length 319;
XX Best Local Similarity 100.0%; Pred. No. 2.4e+02;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 GGCATCGTCAGTTGC 16
XX Db |||||
XX 45 GGCATCGTCAGTTGC 59
XX
XX RESULT 7
XX ADK55125
XX ID ADK55125 standard; DNA; 319 BP.
XX
XX AC ADK55125;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Plant DNA sequence which confers altered metabolic characteristic #2508.
XX
XX KW altered metabolic characteristic; plant; acid metabolism;
XX KW alcohol metabolism; fatty acid metabolism;
XX KW branched fatty acid metabolism; alkaloid metabolism;
XX KW amino acid metabolism; ester metabolism; glyceride metabolism;
XX KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
XX OS Unidentified.
XX
XX PF WO2003020936-A1.
XX
XX PN 13-MAR-2003.
XX
XX PD 30-AUG-2002; 2002WO-US027884.
XX
XX PF 31-AUG-2001; 2001US-0316471P.
XX
XX PA (DOWC ) DOW CHEM CO.
XX
XX PA (DOWC ) DOW AGROSCIENCES LLC.
XX
XX PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX XX WPI; 2003-313091/30.
XX
XX PT Novel genes that confer altered metabolic characteristics in Nicotiana
PI PT benthiana plants, useful for altering the levels of metabolites e.g.
XX XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX PS Claim 1; SEQ ID NO 4352; 2576pp; English.
XX
XX CC The invention comprises DNA sequences which confer an altered metabolic
CC CC characteristic when they are expressed in a plant. The DNA sequences of
CC CC the invention are useful for producing plants with an altered metabolic
CC CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC CC altered glyceride metabolism, altered phenolic metabolism, altered
CC CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC CC invention may be used to provide disease resistance in a plant and gene
CC CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC CC DNA sequence of the invention.
XX
XX SQ Sequence 319 BP; 82 A; 77 C; 93 G; 67 T; 0 U; 0 Other;
XX
XX Query Match 93.8%; Score 15; DB 10; Length 319;
XX Best Local Similarity 100.0%; Pred. No. 2.4e+02;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 GGCATCGTCAGTTGC 16
XX Db |||||
XX 45 GGCATCGTCAGTTGC 59
XX
XX RESULT 8
XX ADK53158
XX ID ADK53158 standard; DNA; 340 BP.
XX
XX AC ADK53158;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Plant DNA sequence which confers altered metabolic characteristic #541.
XX
XX KW altered metabolic characteristic; plant; acid metabolism;
XX KW alcohol metabolism; fatty acid metabolism;
XX KW branched fatty acid metabolism; alkaloid metabolism;
XX KW amino acid metabolism; ester metabolism; glyceride metabolism;
XX KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
XX OS Hypocrea lixii.
XX
XX PF WO2003020936-A1.
XX
XX PN 13-MAR-2003.
XX
XX PD 30-AUG-2002; 2002WO-US027884.
XX
XX PF 31-AUG-2001; 2001US-0316471P.
XX
XX PA (DOWC ) DOW CHEM CO.
XX
XX PA (DOWC ) DOW AGROSCIENCES LLC.
XX
XX PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX XX WPI; 2003-313091/30.
XX
XX PT Novel genes that confer altered metabolic characteristics in Nicotiana
PI PT benthiana plants, useful for altering the levels of metabolites e.g.
XX XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX

```


PS Claim 1; SEQ ID NO 541; 2576pp; English.

XX The invention comprises DNA sequences which confer an altered metabolic

CC characteristic when they are expressed in a plant. The DNA sequences of

CC the invention are useful for producing plants with an altered metabolic

CC characteristic, such as: altered acid metabolism, alcohol metabolism,

CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other

CC base metabolism, altered amino acid metabolism, altered ester metabolism,

CC altered glyceride metabolism, altered phenolic metabolism, altered

CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon

CC metabolism, ketone or quinone metabolism. The DNA sequences of the

CC invention may be used to provide disease resistance in a plant and gene

CC shuffling or sexual PCR procedures. The present nucleic acid represents a

CC DNA sequence of the invention.

XX

XX Sequence 340 BP; 89 A; 81 C; 102 G; 68 T; 0 U; 0 Other;

SEQ

Query Match 93.8%; Score 15; DB 10; Length 340;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16

Db 68 GGCATCGTCAGTTGC 82

RESULT 9

ADK56429

ID ADK56429 standard; DNA; 368 BP.

XX

AC ADK56429;

XX

DT 06-MAY-2004 (first entry)

XX

DE Plant DNA sequence which confers altered metabolic characteristic #3812.

XX

XX altered metabolic characteristic; plant; acid metabolism;

KW alcohol metabolism; fatty acid metabolism;

KW branched fatty acid metabolism; alkaloid metabolism;

KW amino acid metabolism; ester metabolism; glyceride metabolism;

KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;

KW terpene metabolism; isoprenoid metabolism; alkene metabolism;

KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;

KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.

XX

OS Hypocrea lixii.

XX

XX

PN WO2003020936-A1.

XX

XX

PD 13-MAR-2003.

XX

PF 30-AUG-2002; 2002WO-US027884.

XX

XX

PR 31-AUG-2001; 2001US-0316471P.

XX

XX

PA (DOWC) DOW CHEM CO.

PA (DOWC) DOW AGROSCIENCES LLC.

PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;

PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;

DR WPI; 2003-313091/30.

XX

XX Novel genes that confer altered metabolic characteristics in Nicotiana

PT benthamiana plants, useful for altering the levels of metabolites e.g.

PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.

XX

XX Claim 1; SEQ ID NO 3812; 2576pp; English.

PS

XX The invention comprises DNA sequences which confer an altered metabolic

CC characteristic when they are expressed in a plant. The DNA sequences of

CC the invention are useful for producing plants with an altered metabolic

CC characteristic, such as: altered acid metabolism, alcohol metabolism,

CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other

CC base metabolism, altered amino acid metabolism, altered ester metabolism,

CC altered glyceride metabolism, altered phenolic metabolism, altered

CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon

CC metabolism, ketone or quinone metabolism. The DNA sequences of the

CC invention may be used to provide disease resistance in a plant and gene

CC shuffling or sexual PCR procedures. The present nucleic acid represents a

CC DNA sequence of the invention.

XX

XX Sequence 368 BP; 117 A; 82 C; 101 G; 68 T; 0 U; 0 Other;

SEQ

Query Match 93.8%; Score 15; DB 10; Length 368;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16

Db 68 GGCATCGTCAGTTGC 82

RESULT 10

ABN40186/c

ID ABN40186 standard; DNA; 60 BP.

XX

AC ABN40186;

XX

DT 15-JUL-2002 (first entry)

XX

DE Human spliced transcript detection oligonucleotide SEQ ID NO:12934.

XX

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX

OS Homo sapiens.

XX

XX WO200210449-A2.

XX

PD 07-FEB-2002.

XX

XX 20-JUL-2001; 2001WO-1B001903.

XX

XX 28-JUL-2000; 2000US-0221607P.

PR 02-MAY-2001; 2001US-0287724P.

XX

XX (COMP-) COMPUGEN INC.

XX

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX

XX WPI; 2002-257383/30.

DR

XX

XX New oligonucleotide libraries comprising oligonucleotides which

PT selectively hybridize to mRNAs transcribed from a transcription unit of a

PT genome, useful for detecting tissue-, pathology-, and developmental-

PT specific genes.

XX

PS Example 1; SEQ ID NO 12934; 47pp; English.

XX

XX The present invention describes oligonucleotide libraries for detecting

CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-

CC)transcriptome comprises messenger RNAs transcribed from multiple

CC transcription units that populate a genome. The library comprises several

CC oligonucleotides, each capable of hybridising selectively to a set of

CC messenger RNAs transcribed from a given transcription unit of the genome,

CC which encodes one or more messenger RNA splice variants. The

CC oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or

CC quantitatively characterising the corresponding transcriptome, and in

CC detecting RNA transcripts and splice variants of human or animal

CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a particular

CC biological or pathological state, and so allowing the detection of tissue

CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC developmental specific genes; and to detect RNA transcripts and splice
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 60 BP; 20 A; 15 C; 15 G; 10 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 6; Length 60;
 Best Local Similarity 93.8%; Pred. No. 4.2e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 26 CGGCATCGTCAGTTGC 11

RESULT 11
 ACH42666/c
 ID ACH42666 standard; cDNA; 387 BP.

XX ACH42666;

DT 13-OCT-2003 (first entry)

DE Human foetal kidney cDNA #200.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

XX US2003073623-A1.

PN 17-APR-2003.

PD 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

PR (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 29878; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide

CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX SQ Sequence 387 BP; 121 A; 92 C; 97 G; 77 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 9; Length 387;
 Best Local Similarity 93.8%; Pred. No. 5.2e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 270 CGGCATCGTCAGTTGC 255

RESULT 12

ACH32955/c

ID ACH32955 standard; cDNA; 498 BP.

XX ACH32955;

DT 13-OCT-2003 (first entry)

XX Human endothelial cell cDNA #1088.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.

OS Homo sapiens.

XX US2003073623-A1.

PN 17-APR-2003.

PD 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

PR (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 20167; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antibodies specific for it. The present polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 498 BP; 136 A; 124 C; 122 G; 110 T; 0 U; 6 Other;

Query Match 90.0%; Score 14.4; DB 9; Length 498;
Best Local Similarity 93.8%; Pred. No. 5.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
|||||
Db 232 CGGCATCGTCAGTTGC 217

RESULT 13
ID ADB68745/C
XX ADB68745 standard; DNA; 664 BP.
XX
XX ADB68745;
XX AC
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE Pseudomonas chlororaphis phzI DNA.
XX
XX
XX quorum sensing; lux homologue; luxI; ds.
XX
XX Pseudomonas chlororaphis.
XX
XX WO2003057902-A2.
XX
XX 17-JUL-2003.
XX
XX 08-JAN-2003; 2003WO-US000479.
XX
XX 08-JAN-2002; 2002US-0346531P.
XX 07-JAN-2003; 2003US-00338110.
XX
XX (FRAU) FRAUNHOFER USA INC.
XX
XX Fuhrmann JJ, Romesser JA;
XX WPI; 2003-618102/58.
XX
XX Detecting quorum sensing potential of a Gram-negative bacterium in a
XX sample comprises performing a polymerase chain reaction using nucleic
XX acids extracted from a sample containing a microorganism.
XX
XX Disclosure; Fig 10; 86pp; English.
XX
XX The invention relates to a novel method for detecting the quorum sensing
XX potential of a microorganism in a sample which comprises performing PCR
XX using nucleic acids extracted from a sample containing at least one type
XX of microorganism. The method may be useful for detecting the quorum
XX sensing potential of a microorganism in a sample by amplifying a fragment
XX of a lux gene or homologue. The current sequence is that of the luxI
XX homologue DNA of the invention.
XX
SQ Sequence 664 BP; 162 A; 219 C; 167 G; 116 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 10; Length 664;
Best Local Similarity 93.8%; Pred. No. 5.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
|||||
Db 418 CGGCATCGTCAGTTGC 403

RESULT 14
ACA40338
ID ACA40338 standard; DNA; 930 BP.
XX
XX ACA40338;
XX
XX

DT 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #21995.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Mycobacterium tuberculosis.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX P-PSDB; ABU36468.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 28208; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 930 BP; 136 A; 278 C; 355 G; 161 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 8; Length 930;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 828 CGGCATCGGCAGTTGC 843

RESULT 15
ACA38540
ID ACA38540 standard; DNA; 936 BP.
XX AC ACA38540;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #20197.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX OS Mycobacterium bovis.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX DR P-PSDB; ABU34670.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 26410; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 936 BP; 136 A; 281 C; 357 G; 162 T; 0 U; 0 Other;
Query Match 90.0%; Score 14.4; DB 8; Length 936;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 837 CGGCATCGGCAGTTGC 852

RESULT 16
ADS46729/c
ID ADS46729 standard; cDNA; 1056 BP.
XX AC ADS46729;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polynucleotide #1472.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX OS Bacteria.
XX PN US20032333675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 25159; 122pp; English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 1056 BP; 281 A; 188 C; 262 G; 325 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 13; Length 1056;
Best Local Similarity 93.8%; Pred. No. 5.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| |||||||
Db 412 CGGCATCGGCAGTTGC 397

RESULT 17
ABZ13613
ID ABZ13613 standard; DNA; 1137 BP.

AC ABZ13613;

XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1418.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JUN-2001; 2001US-0284647P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 1418; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

Sequence 1137 BP; 310 A; 212 C; 273 G; 342 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 6; Length 1137;
Best Local Similarity 93.8%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| |||||||
Db 533 CGGCATTGTCAGTTGC 548

RESULT 18

AAC33636

ID AAC33636 standard; DNA; 1301 BP.

XX AAC33636;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 3772.

XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126284P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130449P.

XX 28-APR-1999; 99US-0130510P.

XX 30-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 04-MAY-1999; 99US-0132048P.

XX 05-MAY-1999; 99US-0132407P.

XX 06-MAY-1999; 99US-0132484P.

XX 07-MAY-1999; 99US-0132485P.

XX 11-MAY-1999; 99US-0132487P.

XX 14-MAY-1999; 99US-0132487P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134218P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 24-MAY-1999; 99US-0135299P.

XX 25-MAY-1999; 99US-0135353P.

XX 27-MAY-1999; 99US-0136021P.

XX 28-MAY-1999; 99US-0136392P.

XX 01-JUN-1999; 99US-0136782P.

XX 03-JUN-1999; 99US-0137222P.

XX 04-JUN-1999; 99US-0137528P.

XX 07-JUN-1999; 99US-0137502P.

XX 08-JUN-1999; 99US-0137724P.

XX 10-JUN-1999; 99US-0138094P.

XX 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140911P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 27-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159631P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 90.0%; Score 14.4; DB 3; Length 1301;
Best Local Similarity 93.8%; Pred. No. 6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGCATCGTCAGTGC 16
|||||

Db 574 CGGCATTGTCAGTTGC 589

RESULT 19
ADT46111
ID ADT46111 standard; cDNA; 1362 BP.
XX AC
XX AC ADT46111;
DT 02-DEC-2004 (first entry)
XX DE
XX DE Bacterial polynucleotide #20862.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.
XX OS Bacteria.
XX OS US2003233675-A1.
XX PN
XX XX
XX 18-DEC-2003.
XX PD
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX DR
XX XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX PT
XX PS Claim 1; SEQ ID NO 44549; 122pp; English.
XX XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX XX
XX Sequence 1362 BP; 242 A; 464 C; 423 G; 233 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 13; Length 1362;
Best Local Similarity 93.8%; Pred. No. 6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
| | | | | | | | | | | | | | | | | |
DB 417 CGGCATCGTCAGTTGC 432

RESULT 20
ACA36906/C
ID ACA36906 standard; DNA; 1392 BP.
XX AC
XX AC ACA36906;
XX DT 19-JUN-2003 (first entry)
XX XX
XX DE Prokaryotic essential gene #18563.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX OS Listeria monocytogenes.
XX XX
XX WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX DR P-PSDB; ABUS3036.
XX PT
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 24776; 1766pp; English.
XX XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC of a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 1392 BP; 343 A; 214 C; 336 G; 499 T; 0 U; 0 Other;
 Query Match 90.0%; Score 14.4; DB 8; Length 1392;
 Best Local Similarity 93.8%; Pred. No. 6e+02; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 DB 762 CGGCATCGTCGGTTGC 747

RESULT 21
 ACF70905/c
 ID ACF70905 standard; DNA; 1563 BP.
 AC ACF70905;
 XX
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens nucleotide sequence #9372.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 XX
 FN WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 XX Claim 2; SEQ ID NO 9372; 1205pp; French.
 PS
 XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of *P. luminescens*
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than *P. luminescens* and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by *P.*
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The

CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which *P.*
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated *P. luminescens* genes

XX
 SQ Sequence 1563 BP; 387 A; 363 C; 509 G; 304 T; 0 U; 0 Other;
 Query Match 90.0%; Score 14.4; DB 10; Length 1563;
 Best Local Similarity 93.8%; Pred. No. 6.1e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 DB 882 CGGCATCGTCAGTAGC 867

RESULT 22
 AAQ96251/c
 ID AAQ96251 standard; DNA; 1632 BP.
 XX
 AC AAQ96251;
 XX
 DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-DEC-1995 (first entry)
 XX
 DE Fructosyltransferase gene of *A. diazotrophicus*.
 XX
 KW Fructosyltransferase; FTase; fructooligosaccharide; sweetener; fructan;
 KW ds.
 XX
 OS Gluconacetobacter diazotrophicus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1632
 FT /*tag= a
 XX
 FN EP663442-A1.
 XX
 PD 19-JUL-1995.
 XX
 PF 22-DEC-1994; 94EP-00203737.
 XX
 PR 23-DEC-1993; 93CU-00000125.
 XX
 PA (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
 XX
 XX Sosa JGA, Garcia LH, Gonzalez AC, Sosa GS;
 PI WPI; 1995-247529/33.
 DR P-PSDB; AAR79142.
 XX
 XX New fructosyltransferase from Acetobacter diazotrophicus - for the
 PT prodn. of fructo-oligosaccharide and fructan cpds. from sucrose, useful
 FT e.g. as low calorie sweeteners.
 XX
 XX Claim 1; Page 10-11; 16pp; English.
 PS
 XX A genomic library of *A. diazotrophicus* SRT4 was produced in pPW12 and
 CC transformed to a lev-neg. mutant of the same strain. Colonies that
 CC recovered the mucous phenotype were selected; 2 recombinant cosmids each
 CC contained the same 7.8 kb insert, and the Frase gene was localized to a
 CC 2.3 kb fragment that was cloned into pUC18 to give pUC18S23; the sequence
 CC of the Frase gene in this plasmid is given in AAQ96251. Cloning allows
 CC expression of recombinant FTase in *E. coli* and *Pichia pastoris*. (Updated
 CC on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to
 CC standardise OS field)

XX
 SQ Sequence 1632 BP; 307 A; 551 C; 485 G; 289 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 2; Length 1632;
Best Local Similarity 93.8%; Pred. No. 6.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
| | | | | | | | | | | | | | | | | | | | | |
Db 231 CGGCATCGTCAGTTGC 216

RESULT 23
AAC77952/c
ID AAC77952 standard; cDNA; 1714 BP.

XX AC AAC77952;
XX DE 08-FEB-2001 (first entry)
XX DE Human cancer associated gene sequence SEQ ID NO:346.
XX KW Human; cancer associated gene; cancer antigen; detection; cancer;
XX KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
XX KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
XX KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
XX KW dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
XX KW vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;
XX KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX KW allergic reaction; graft versus host disease; organ rejection;
XX KW haemostatic; thrombolytic; cardiovascular disorder; infection;
XX KW neurological disease; drug screening; ss.
XX OS Homo sapiens.
XX FN WO200055350-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005882.
XX PR 12-MAR-1999; 99US-0124270P.
XX XX (HUMA-) HUMAN GENOME SCI INC.
XX PA Rosen CA, Ruben SM;
XX PI WPI; 2000-587533/55.
XX DR P-ESDB; AAB43743.
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX PS Claim 1; Page 896-897; 2352pp; English.

XX AC AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerable; immunomodulator;
XX antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
XX neotropic; vasotropic; antipsoriatic and angiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention

XX SQ Sequence 1714 BP; 539 A; 332 C; 409 G; 433 T; 0 U; 1 Other;
Query Match 90.0%; Score 14.4; DB 3; Length 1714;
Best Local Similarity 93.8%; Pred. No. 6.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
| | | | | | | | | | | | | | | | | | | | | |
Db 1659 CGGCATCGTCAGTTGC 1644

RESULT 24
ADS47843/c
ID ADS47843 standard; cDNA; 1734 BP.

XX AC ADS47843;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polynucleotide #2586.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polynucleotide; gene; ss.
XX OS Bacteria.
XX FN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX PT New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 26273; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1734 BP; 538 A; 334 C; 364 G; 498 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 13; Length 1734;
Best Local Similarity 93.8%; Pred. No. 6.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 1429 CGGCACGTCAGTTGC 1414

RESULT 25
AAD57635/c
ID AAD57635 standard; DNA; 1797 BP.

XX AAD57635;

DT 20-NOV-2003 (first entry)

XX Rice disease resistance gene, PIP1C1.

XX Rice; abiotic stress tolerance; pathogen resistance; disease resistance;
XX grain quality; nutritional content; plant yield; PIP1C1; plant; gene; ds.

XX Oryza sativa.

XX Key Location/Qualifiers
FH 1..1797
CDS /*tag= a
FT /product= "Rice PIP1C1 protein"

XX WO2003048319-A2.

XX 12-JUN-2003.

XX 27-NOV-2002; 2002WO-US038359.

XX 30-NOV-2001; 2001US-0334501P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Sainz MB, Salmeron J, Weislo L;

XX WPI; 2003-505288/47.

XX P-PSDB; AAE38267.

XX New nucleic acid from Oryza sativa, useful for altering abiotic stress
XX tolerance, pathogen or disease resistance or the grain quality,
XX nutritional content or yield in a plant.

XX Claim 2; Page 139-140; 223pp; English.

XX The invention relates to nucleic acid molecules from rice encoding
XX proteins for abiotic stress tolerance, enhanced pathogen or disease
XX resistance and altered nutritional quality. The sequences of the
XX invention are useful for altering abiotic stress tolerance, pathogen or
XX disease resistance or the grain quality, nutritional content or yield in
XX a plant. The present sequence is rice disease resistance gene, PIP1C1

XX Sequence 1797 BP; 447 A; 541 C; 470 G; 339 T; 0 U; 0 Other;

Query Match 90.0%; Score 14.4; DB 10; Length 1797;
Best Local Similarity 93.8%; Pred. No. 6.2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 859 CGGCATCGTCAGTTGC 844

Search completed: March 11, 2005, 04:20:02
Job time : 25.4069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 6.48101 Seconds
(without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-19

Perfect score: 16

Sequence: 1 cggcatcgctcagttgc 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14.4	90.0	1632	1	US-08-362-232-1
C 2	14.4	90.0	1632	1	US-08-814-196-1
C 3	14.4	90.0	2328	4	US-09-252-931A-5729
C 4	14.4	90.0	2511	4	US-09-023-655-935
C 5	14.4	90.0	3099	4	US-09-902-540-5065
C 6	14.4	90.0	4230	4	US-09-252-991A-5711
C 7	14.4	90.0	4941	4	US-09-252-991A-5741
C 8	14.4	90.0	28194	4	US-08-902-540-1250
C 9	14.4	90.0	4403765	3	US-09-103-840A-2
C 10	14.4	90.0	4411529	3	US-09-103-840A-1
C 11	14	87.5	206433	4	US-09-949-016-13527
C 12	14	87.5	254778	4	US-09-949-016-12417
C 13	13.4	83.8	299	3	US-09-118-554-42
C 14	13.4	83.8	299	3	US-09-118-627-42
C 15	13.4	83.8	299	3	US-09-602-877A-42
C 16	13.4	83.8	354	4	US-09-489-039A-1808
C 17	13.4	83.8	601	4	US-09-949-016-108325
C 18	13.4	83.8	601	4	US-09-949-016-175637
C 19	13.4	83.8	864	4	US-09-949-016-4991
C 20	13.4	83.8	1237	4	US-09-270-767-12517
C 21	13.4	83.8	1299	1	US-08-453-472-3
C 22	13.4	83.8	1299	1	US-08-038-948-3
C 23	13.4	83.8	1299	1	US-08-453-952-3
C 24	13.4	83.8	1299	2	US-08-862-903-3
C 25	13.4	83.8	1299	2	US-08-484-158B-60
C 26	13.4	83.8	1326	4	US-09-902-540-5599
C 27	13.4	83.8	1422	4	US-09-724-623-42

28	13.4	83.8	1442	4	US-09-634-238-82	Sequence 82, Appl
29	13.4	83.8	1500	4	US-09-603-208A-233	Sequence 233, App
C 30	13.4	83.8	1530	4	US-09-489-039A-3179	Sequence 3179, Ap
C 31	13.4	83.8	1587	4	US-09-489-039A-6670	Sequence 6670, Ap
C 32	13.4	83.8	2061	3	US-09-008-271A-16	Sequence 16, Appl
C 33	13.4	83.8	2061	3	US-09-705-448-2	Sequence 2, Appl
C 34	13.4	83.8	2115	4	US-08-388-852B-1	Sequence 1, Appl
C 35	13.4	83.8	2130	3	US-09-056-105-1	Sequence 1, Appl
C 36	13.4	83.8	2131	4	US-08-234-784B-91	Sequence 91, Appl
C 37	13.4	83.8	2172	2	US-08-417-174-26	Sequence 26, Appl
C 38	13.4	83.8	2172	2	US-08-231-565A-26	Sequence 26, Appl
C 39	13.4	83.8	2172	2	US-09-007-961-26	Sequence 26, Appl
C 40	13.4	83.8	2172	3	US-09-267-439-26	Sequence 26, Appl
C 41	13.4	83.8	2172	4	US-09-073-138-26	Sequence 26, Appl
C 42	13.4	83.8	2469	4	US-09-489-039A-6749	Sequence 6749, Ap
C 43	13.4	83.8	5281	4	US-09-949-016-824	Sequence 824, App
C 44	13.4	83.8	6380	4	US-09-902-540-891	Sequence 891, App
C 45	13.4	83.8	7791	4	US-09-949-016-13900	Sequence 13900, A

ALIGNMENTS

RESULT 1
US-08-362-232-1/c
; Sequence 1, Application US/08362232
; Patent No. 5641667
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Fructosyltransferase Enzyme, Method
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.0 for DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,232
; FILING DATE: 22-December-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CU 125/93
; FILING DATE: 23-December-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Barton, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1632 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-362-232-1

Query Match 90.0%; Score 14.4; DB 1; Length 1632;

Best Local Similarity 93.8%; Pred. No. 1.3e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| |||||

Db 231 CGGCATCGTCAGTTGC 216

RESULT 2

US-08-814-196-1/c

; Sequence 1, Application US/08814196

; Patent No. 5731173

; GENERAL INFORMATION:

; APPLICANT: Fructosyltransferase Enzyme, Method

; TITLE OF INVENTION: For its Production and DNA Encoding the Enzyme.

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann & Baron

; STREET: 350 Jericho Turnpike

; CITY: Jericho

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 Mb

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 6.0 for DOS

CURRENT APPLICATION DATA: US/08/814,196

APPLICATION NUMBER: US/08/814,196

FILING DATE: 10-MAR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/362,232

FILING DATE: 22-December-1994

APPLICATION NUMBER: CU 125/93

FILING DATE: 23-December-1993

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-29

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1632 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

US-08-814-196-1

Query Match 90.0%; Score 14.4; DB 1; Length 1632;

Best Local Similarity 93.8%; Pred. No. 1.3e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 231 CGGCATCGTCAGTTGC 216

RESULT 3

US-09-252-991A-5729

; Sequence 5729, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 5729

; LENGTH: 2328

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-5729

Query Match 90.0%; Score 14.4; DB 4; Length 2328;

Best Local Similarity 93.8%; Pred. No. 1.3e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 289 CGGCATCGTCAGTTGC 304

RESULT 4

US-09-023-655-935/c

; Sequence 935, Application US/09023655

; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; TITLE OF INVENTION: EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 935:

SEQUENCE CHARACTERISTICS:

LENGTH: 2511 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g1381141

US-09-023-655-935

Query Match 90.0%; Score 14.4; DB 4; Length 2511;

Best Local Similarity 93.8%; Pred. No. 1.3e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 1505 CGGCATCGTCAGTTGC 1490

```

; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5741
; LENGTH: 4941
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5741

Query Match          90.0%; Score 14.4; DB 4; Length 4941;
Best Local Similarity 93.8%; Pred. No. 1.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
    |||||
Db 1551 CGGCATCGTCAGTTTC 1566

RESULT 8
US-09-902-540-1250
; Sequence 1250, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1250
; LENGTH: 28194
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1250

Query Match          90.0%; Score 14.4; DB 4; Length 28194;
Best Local Similarity 93.8%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
    |||||
Db 4908 CGGCATCGTCAGTCGC 4923

RESULT 9
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis

```

```

; CURRENT APPLICATION NUMBER: US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5065
; LENGTH: 3099
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5065

Query Match          90.0%; Score 14.4; DB 4; Length 3099;
Best Local Similarity 93.8%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
    |||||
Db 321 CGGCATCGTCAGTCGC 336

RESULT 6
US-09-252-991A-5711/c
; Sequence 5711, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5711
; LENGTH: 4230
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5711

Query Match          90.0%; Score 14.4; DB 4; Length 4230;
Best Local Similarity 93.8%; Pred. No. 1.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
    |||||
Db 2299 CGGCATCGTCAGTTTC 2284

RESULT 7
US-09-252-991A-5741
; Sequence 5741, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

```

```
;
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          90.0%; Score 14.4; DB 3; Length 4403765;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 3238851 CGGCATCGTCAGTTGC 3238866
|||||

RESULT 10
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match          90.0%; Score 14.4; DB 3; Length 4411529;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 3244575 CGGCATCGTCAGTTGC 3244590
|||||

RESULT 11
US-09-949-016-13527/c
; Sequence 13527, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13527
; LENGTH: 206433
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13527

Query Match          87.5%; Score 14; DB 4; Length 206433;
```

```
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTT 14
Db 163150 CGGCATCGTCAGTT 163137
|||||

RESULT 12
US-09-949-016-12417/c
; Sequence 12417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12417
; LENGTH: 254778
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12417
```

```
Query Match          87.5%; Score 14; DB 4; Length 254778;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CGGCATCGTCAGTT 14
Db 163495 CGGCATCGTCAGTT 163482
|||||
```

```
RESULT 13
US-09-118-554-42/c
; Sequence 42, Application US/09118554A
; Patent No. 6365348
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 21021.450C1
; CURRENT APPLICATION NUMBER: US/09/118,554A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 08/988,255
; EARLIER FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (228)
; OTHER INFORMATION: Wherein n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (282)
; OTHER INFORMATION: Wherein n is a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (298)
```

OTHER INFORMATION: wherein is a, c, g or t
US-09-118-554-42

Query Match 83.8%; Score 13.4; DB 3; Length 299;
Best Local Similarity 93.3%; Pred. No. 4e+02;
Matches 14; Conservative 0; Mismatches 1; Indels

Qy 2 GGCATCGTCA GTTC 16
Db 54 GGCAACGTCA GTTC 40

```

RESULT 14
US-09-118-627-42/c
; Sequence 42, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
; TITLE OF INVENTION: AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C1
; CURRENT APPLICATION NUMBER: US/09/118.627A
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(299)
; OTHER INFORMATION: n = A,T,C or G
US-09-118-627-42

```

Query Match 83.8%; Score 13.4; DB 3; Length 299;
Best Local Similarity 93.3%; Pred. No. 4e+02;
Matches 14; Conservative 0; Mismatches 1; Indels

Qy 2 GGCATCGTCAGTTGC 16
Db 54 GGCAACGTCAGTTGC 40

```

RESULT 15
US-09-602-877A-42/c
; Sequence 42, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(299)
; OTHER INFORMATION: n = A,T,C or G
US-09-602-877A-42

```

Query Match 83.8%; Score 13.4; DB 3; Length 299;
Best Local Similarity 93.3%; Pred. No. 4e+02;
Matches 14; Conservative 0; Mismatches 1; Indels

```

Qy      2  GGCATCGTCAGTTGC  16
      |||||
Db      54  GGCACGTCAGTTGC  40

```

Query Match 83.8%; Score 13.4; DB 4; Length 354;
Best Local Similarity 93.3%; Pred. No. 4e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16
Db 257 GGCATCGTCAATTGC 271

```

RESULT 17
US-09-949-016-108325
; Sequence 108325, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108325
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-108325

```

```
Query Match      83.8%; Score 13.4; DB 4; Length 601;
Best Local Similarity 93.3%; Pred. No. 4.3e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 CGGCATCGTCAGTTG 15
Db 350 CAGCATCGTCAGTTG 364

RESULT 18
US-09-949-016-175637/c
; Sequence 175637, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 175637
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-175637

Query Match 83.8%; Score 13.4; DB 4; Length 601;
 Best Local Similarity 93.3%; Pred. No. 4.3e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCATCGTCAGTTGC 16
 |||||
 Db 380 GGCATCGTCAGTTAC 366

RESULT 19
 US-09-949-016-4991/c
 ; Sequence 4991, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4991
 ; LENGTH: 864
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-4991

Query Match 83.8%; Score 13.4; DB 4; Length 864;
 Best Local Similarity 93.3%; Pred. No. 4.4e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGCATCGTCAGTTGC 16
 |||||
 Db 357 GGCATCGTCAGTTAC 343

RESULT 20
 US-09-270-767-12517/c
 ; Sequence 12517, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12517
 ; LENGTH: 1237
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-12517

Query Match 83.8%; Score 13.4; DB 4; Length 1237;
 Best Local Similarity 93.3%; Pred. No. 4.6e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTG 15
 |||||
 Db 961 CGGCATCGTCAGTCG 947

RESULT 21
 US-08-453-472-3/c
 ; Sequence 3, Application US/08453472
 ; Patent No. 5628846
 ; GENERAL INFORMATION:
 ; APPLICANT: DEAN, JURRIEN
 ; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
 ; BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
 ; TITLE OF INVENTION: POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN
 ; STREET: 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/453,472
 ; FILING DATE: 30-May-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/038,948
 ; FILING DATE: 26-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/930,462
 ; FILING DATE: 20-AUG-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/364,379
 ; FILING DATE: 12-JUN-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DOROTHY R. AULT
 ; REGISTRATION NUMBER: 36,434
 ; REFERENCE/DOCKET NUMBER: 2026-4032 US3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; TELEX: 421792
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1299
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cdna
 ; ORIGINAL SOURCE:
 ; ORGANISM: human
 ; STRAIN:
 ; INDIVIDUAL ISOLATE:
 ; DEVELOPMENTAL STAGE:

;
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:
; NAME/KEY: ZP3
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP3 cDNA
; US-08-453-472-3

Query Match 83.8%; Score 13.4; DB 1; Length 1299;
Best Local Similarity 93.3%; Pred. No. 4.6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GGCACTCGTCAGTTGC 16
| | | | | | | | | | | | | | | |
Db 337 GGCACTCGTCAGTTAC 323

RESULT 22

US-08-038-948-3/c

; Sequence 3, Application US/08038948

; Patent No. 5641487

; GENERAL INFORMATION:

; APPLICANT: DEAN, JURRIEN

; TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CUSHMAN, DABY & CUSHMAN

; STREET: 1100 New York Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.

; ZIP: 20005-3918

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/038,948

; FILING DATE: 26-MAR-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/930,462

; FILING DATE: 20-AUG-1992

; APPLICATION DATA:

; APPLICATION NUMBER: US 07/364,379

; FILING DATE: 12-JUN-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: SCOTT, Watson T.

; REGISTRATION NUMBER: 26,581

; REFERENCE/DOCKET NUMBER: 99152/E-266-88/2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 861-3000

; TELEFAX: (202) 822-0944

; TELEX: 6714627 CUSH

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1299 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-038-948-3

Query Match 83.8%; Score 13.4; DB 1; Length 1299;

Best Local Similarity 93.3%; Pred. No. 4.6e+02;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 GGCACTCGTCAGTTGC 16
| | | | | | | | | | | | | | | |
Db 337 GGCACTCGTCAGTTAC 323

RESULT 23

US-08-453-952-3/c

; Sequence 3, Application US/08453952

; Patent No. 5672488

; GENERAL INFORMATION:

; APPLICANT: DEAN, JURRIEN

; TITLE OF INVENTION: CONTRACEPTIVE VACCINE

; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/453,952

; FILING DATE: 30-May-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/038,948

; FILING DATE: 26-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/930,462

; FILING DATE: 20-AUG-1992

; APPLICATION DATA:

; APPLICATION NUMBER: US 07/364,379

; FILING DATE: 12-JUN-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: DOROTHY R. AUTH

; REGISTRATION NUMBER: 36,434

; REFERENCE/DOCKET NUMBER: 2026-4032 US4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1299

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; ORIGINAL SOURCE:

; ORGANISM: human

; STRAIN:

; INDIVIDUAL ISOLATE:

; DEVELOPMENTAL STAGE:

; HAPLOTYPE:

; TISSUE TYPE:

; CELL TYPE:

; CELL LINE:

; ORGANELLE:

; FEATURE:

; NAME/KEY: ZP3

; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: human ZP3 cDNA

US-08-453-952-3

Query Match 83.8%; Score 13.4; DB 1; Length 1299;

```
Best Local Similarity 93.3%; Pred. No. 4.6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16
Db 337 GGCATCGTCAGTTAC 323

RESULT 24
US-08-862-903-3/c
; Sequence 3, Application US/08862903
; Patent No. 5916768
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,903
; FILING DATE: 30-May-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; NAME: DOROTHY R. AUTH
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELL:
; FEATURE:
; NAME/KEY: ZP3
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: human ZP3 cdna

US-08-862-903-3
Query Match 83.8%; Score 13.4; DB 2; Length 1299;
Best Local Similarity 93.3%; Pred. No. 4.6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16
Db 337 GGCATCGTCAGTTAC 323

RESULT 25
US-08-484-158B-60/c
; Sequence 60, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; TITLE OF INVENTION: Immunocontraception
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "human ZPC"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1282
; US-08-484-158B-60

Query Match 83.8%; Score 13.4; DB 2; Length 1299;
Best Local Similarity 93.3%; Pred. No. 4.6e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 2 GGCATCGTCAGTTGC 16
|||
Db 337 GGCATCGTCAGTTAC 323

Search completed: March 11, 2005, 13:13:21
Job time : 16.481 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 79.9325 Seconds
(without alignment)
1190.710 Million cell updates/sec

Title: US-09-674-277-19

Perfect score: 16

Sequence: 1 cggcatcgctcagttgc 16

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	14.4	90.0	25	19	US-10-719-900-375244	
c	2	14.4	90.0	60	10	US-09-908-975-12934
c	3	14.4	90.0	182	18	US-10-425-115-41603
c	4	14.4	90.0	387	10	US-09-918-995-29878
c	5	14.4	90.0	414	18	US-10-437-963-83252
c	6	14.4	90.0	498	10	US-09-918-995-20167
c	7	14.4	90.0	664	17	US-10-338-110-22
c	8	14.4	90.0	767	18	US-10-425-115-101173
c	9	14.4	90.0	930	17	US-10-282-122A-28208
c	10	14.4	90.0	936	17	US-10-282-122A-26410
c	11	14.4	90.0	985	17	US-10-425-114-6662
						Sequence 375244, A
						Sequence 12934, A
						Sequence 41603, A
						Sequence 29878, A
						Sequence 83252, A
						Sequence 20167, A
						Sequence 22, App1
						Sequence 101173, A
						Sequence 28208, A
						Sequence 26410, A
						Sequence 6662, App1

c	12	14.4	90.0	1056	17	US-10-369-493-25159
c	13	14.4	90.0	1137	9	US-09-938-842A-1418
c	14	14.4	90.0	1137	11	US-09-938-842A-1418
c	15	14.4	90.0	1332	17	US-10-424-599-13630
c	16	14.4	90.0	1349	18	US-10-437-963-41186
c	17	14.4	90.0	1362	17	US-10-369-493-44549
c	18	14.4	90.0	1392	17	US-10-282-122A-24776
c	19	14.4	90.0	1608	18	US-10-437-963-43552
c	20	14.4	90.0	1714	9	US-09-925-301-346
c	21	14.4	90.0	1734	17	US-10-369-493-26273
c	22	14.4	90.0	1797	18	US-10-491-733-23
c	23	14.4	90.0	1848	18	US-10-437-963-7327
c	24	14.4	90.0	1878	18	US-10-437-963-41189
c	25	14.4	90.0	1945	18	US-10-425-115-101174
c	26	14.4	90.0	2090	17	US-10-425-114-13165
c	27	14.4	90.0	2209	18	US-10-437-963-7325
c	28	14.4	90.0	2511	17	US-10-439-703-64
c	29	14.4	90.0	2511	17	US-10-641-643-935
c	30	14	87.5	695	18	US-10-425-115-143802
c	31	14	87.5	774	9	US-09-738-626-320
c	32	14	87.5	1285	17	US-10-282-122A-36930
c	33	14	87.5	1461	17	US-10-282-122A-14037
c	34	14	87.5	1551	17	US-10-282-122A-40133
c	35	14	87.5	3309400	9	US-09-738-626-1
c	36	13.4	83.8	60	10	US-09-908-975-15299
c	37	13.4	83.8	132	9	US-09-864-761-18565
c	38	13.4	83.8	235	17	US-10-424-599-67863
c	39	13.4	83.8	248	18	US-10-425-115-12917
c	40	13.4	83.8	269	18	US-10-425-115-96142
c	41	13.4	83.8	275	18	US-10-425-115-38274
c	42	13.4	83.8	291	18	US-10-437-963-78465
c	43	13.4	83.8	299	9	US-09-745-288-42
c	44	13.4	83.8	299	17	US-10-453-919-42
c	45	13.4	83.8	354	18	US-10-425-115-144416

ALIGNMENTS

RESULT 1

US-10-719-900-375244
; Sequence 375244, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 375244
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-375244

Query Match 90.0%; Score 14.4; DB 19; Length 25;
Best Local Similarity 93.8%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 4 CGGCATCGTCAGTTCC 19

RESULT 2

US-09-908-975-12934/c
; Sequence 12934, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:

```
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Eliat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12934
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-908-975-12934

Query Match          90.0%; Score 14.4; DB 10; Length 60;
Best Local Similarity 93.8%; Pred. No. 6.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 26 CGGCATCGTCAGTTGC 11

RESULT 3
US-10-425-115-41603/c
; Sequence 41603, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 41603
; LENGTH: 182
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_137943C.1
US-10-425-115-41603

Query Match          90.0%; Score 14.4; DB 18; Length 182;
Best Local Similarity 93.8%; Pred. No. 6.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 97 CGGCATCATCAGTTGC 82

RESULT 4
US-09-918-995-29878/c
; Sequence 29878, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: FROM VARIOUS CDNA LIBRARIES
; CURRENT FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: US/09/918,995
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
```

```
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29878
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-29878

Query Match          90.0%; Score 14.4; DB 10; Length 387;
Best Local Similarity 93.8%; Pred. No. 6.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 270 CGGCATCGTCAGTTGC 255

RESULT 5
US-10-437-963-83252/c
; Sequence 83252, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83252
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82601C.1
US-10-437-963-83252

Query Match          90.0%; Score 14.4; DB 18; Length 414;
Best Local Similarity 93.8%; Pred. No. 6.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
Db 393 CAGCATCGTCAGTTGC 378

RESULT 6
US-09-918-995-20167/c
; Sequence 20167, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: FROM VARIOUS CDNA LIBRARIES
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20167
; LENGTH: 498
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(498)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-20167

Query Match          90.0%; Score 14.4; DB 10; Length 498;
Best Local Similarity 93.8%; Pred. No. 6.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCAGTTGC 16
Db      232 CGGCATCGTCAGTTGC 217

RESULT 7
US-10-338-110-22/c
; Sequence 22, Application US/10338110
; Publication No. US20040023254A1
; GENERAL INFORMATION:
; APPLICANT: Fuhrmann, Jeffry J.
; APPLICANT: Romesser, James A.
; TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial
; FILE REFERENCE: HER-0056
; CURRENT APPLICATION NUMBER: US/10/338,110
; CURRENT FILING DATE: 2003-01-07
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Pseudomonas aureofaciens
US-10-338-110-22

Query Match          90.0%; Score 14.4; DB 17; Length 664;
Best Local Similarity 93.8%; Pred. No. 6.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCAGTTGC 16
Db      418 CGGCATCGCCAGTTGC 403

RESULT 8
US-10-425-115-101173/c
; Sequence 101173, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101173
; LENGTH: 767
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(767)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_23783C.1
US-10-425-115-101173

Query Match          90.0%; Score 14.4; DB 10; Length 498;
Best Local Similarity 93.8%; Pred. No. 6.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCAGTTGC 16
Db      232 CGGCATCGTCAGTTGC 217

RESULT 7
US-10-338-110-22/c
; Sequence 22, Application US/10338110
; Publication No. US20040023254A1
; GENERAL INFORMATION:
; APPLICANT: Fuhrmann, Jeffry J.
; APPLICANT: Romesser, James A.
; TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial
; FILE REFERENCE: HER-0056
; CURRENT APPLICATION NUMBER: US/10/338,110
; CURRENT FILING DATE: 2003-01-07
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Pseudomonas aureofaciens
US-10-338-110-22

Query Match          90.0%; Score 14.4; DB 17; Length 664;
Best Local Similarity 93.8%; Pred. No. 6.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCAGTTGC 16
Db      418 CGGCATCGCCAGTTGC 403

RESULT 8
US-10-425-115-101173/c
; Sequence 101173, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101173
; LENGTH: 767
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(767)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_23783C.1
US-10-425-115-101173

Query Match          90.0%; Score 14.4; DB 18; Length 767;
Best Local Similarity 93.8%; Pred. No. 6.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCAGTTGC 16
Db      362 CGGCATCGTCAGTTGC 347

RESULT 9
US-10-282-122A-28208
; Sequence 28208, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28208
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28208

Query Match          90.0%; Score 14.4; DB 17; Length 930;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCAGTTGC 16
Db      828 CGGCATCGCAGTTGC 843

RESULT 10
US-10-282-122A-26410
; Sequence 26410, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26410
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Mycobacterium bovis
US-10-282-122A-26410

Query Match          90.0%; Score 14.4; DB 17; Length 936;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CGGCATCGTCAGTTGC 16
      |||||
DB      837 CGGCATCGGCAGTTGC 852

RESULT 11
US-10-425-114-6662
; Sequence 6662, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6662
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26410
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Mycobacterium bovis
US-10-282-122A-26410

Query Match          90.0%; Score 14.4; DB 17; Length 936;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 CGGCATCGTCAGTTGC 16
      |||||
DB      837 CGGCATCGGCAGTTGC 852

RESULT 11
US-10-425-114-6662
; Sequence 6662, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6662
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

```


US-09-938-842A-1418

Query Match 90.0%; Score 14.4; DB 9; Length 1137;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
||||| |||||||
DB 533 CGGCATTGTCAGTTGC 548

RESULT 14

US-09-938-842A-1418
; Sequence 1418, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1418
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1418

Query Match 90.0%; Score 14.4; DB 11; Length 1137;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
||||| |||||||
DB 533 CGGCATTGTCAGTTGC 548

RESULT 15

US-10-424-599-13630/c
; Sequence 13630, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 13630
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(1332)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_112315C.1
US-10-424-599-13630

Query Match 90.0%; Score 14.4; DB 17; Length 1332;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
||||| |||||||
DB 851 CGGCCTCGTCAGTTGC 836

RESULT 16

US-10-437-963-41186
; Sequence 41186, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41186
; LENGTH: 1349
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44555C.1
US-10-437-963-41186

Query Match 90.0%; Score 14.4; DB 18; Length 1349;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
||||| |||||||
DB 551 CGCATCGTCAGTTGC 566

RESULT 17

US-10-369-493-44549
; Sequence 44549, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44549
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-44549

Query Match 90.0%; Score 14.4; DB 17; Length 1362;
Best Local Similarity 93.8%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 417 CGGCATCGTCAATGTC 432

RESULT 18
 US-10-282-122A-24776/c
 ; Sequence 24776, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zvekind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: EUIPA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282.122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-07-27
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR FILING DATE: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR FILING DATE: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR FILING DATE: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR FILING DATE: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR FILING DATE: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 24776
 ; LENGTH: 1392
 ; TYPE: DNA
 ; ORGANISM: Listeria monocytogenes
 US-10-282-122A-24776

Query Match 90.0%; Score 14.4; DB 17; Length 1392;
 Best Local Similarity 93.8%; Pred. No. 6.6e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 762 CGGCATCGTCCGTTGC 747

RESULT 19
 US-10-437-963-43652/c
 ; Sequence 43652, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Bouharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 43652
 ; LENGTH: 1608
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_46789C.1
 US-10-437-963-43652

Query Match 90.0%; Score 14.4; DB 18; Length 1608;
 Best Local Similarity 93.8%; Pred. No. 6.7e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 948 CAGCATCGTCAGTTGC 933

RESULT 20
 US-09-925-301-346/c
 ; Sequence 346, Application US/09925301
 ; Patent No. US20020052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR FILING DATE: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1694
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 346
 ; LENGTH: 1714
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (21)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-925-301-346

Query Match 90.0%; Score 14.4; DB 9; Length 1714;
 Best Local Similarity 93.8%; Pred. No. 6.7e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 1659 CGGCATCGTCACTTGC 1644

RESULT 21
 US-10-369-493-26273/c
 ; Sequence 26273, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng

```
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26273
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26273

Query Match          90.0%; Score 14.4; DB 17; Length 1734;
Best Local Similarity 93.8%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| ||||| ||||| ||||| |||||
Db 1429 CGGCACGTCAGTTGC 1414

RESULT 22
US-10-491-733-23/c
; Sequence 23, Application US/10491733
; Publication No. US20040219675A1
; GENERAL INFORMATION:
; APPLICANT: Syngenta Participations AG
; APPLICANT: Sainz, Manuel
; APPLICANT: Salmerton, John
; APPLICANT: Weislo, Laura J.
; TITLE OF INVENTION: Nucleic Acid Molecules from Rice Encoding Proteins for Abiotic St
; TITLE OF INVENTION: Resistance, Yield, Disease Resistance and Nutritional Quality an
; FILE REFERENCE: 60127WOPCT
; CURRENT APPLICATION NUMBER: US/10/491,733
; CURRENT FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: 60/334,501
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-491-733-23

Query Match          90.0%; Score 14.4; DB 18; Length 1797;
Best Local Similarity 93.8%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| ||||| ||||| ||||| |||||
Db 859 CGGCATCGTCAGTTGC 844

RESULT 23
US-10-437-963-7327/c
; Sequence 7327, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
```

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 7327
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13933C.1
US-10-437-963-7327

Query Match          90.0%; Score 14.4; DB 18; Length 1848;
Best Local Similarity 93.8%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| ||||| ||||| ||||| |||||
Db 859 CGGCATCGTCAGTTGC 844

RESULT 24
US-10-437-963-41189
; Sequence 41189, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41189
; LENGTH: 1878
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44558C.1
US-10-437-963-41189

Query Match          90.0%; Score 14.4; DB 18; Length 1878;
Best Local Similarity 93.8%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
||||| ||||| ||||| ||||| |||||
Db 453 CGACATCGTCAGTTGC 468

RESULT 25
US-10-425-115-101174
; Sequence 101174, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101174
; LENGTH: 1945
; TYPE: DNA
```

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1945)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_23784C.1
US-10-425-115-101174

Query Match          90.0%; Score 14.4; DB 18; Length 1945;
Best Local Similarity 93.8%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGCATCGTCACATTGC 16
        |||||
Db      392 CGGCATCGTCGGTTGC 407
```

Search completed: March 12, 2005, 00:25:26
Job time : 81.9325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 211.723 Seconds
(without alignments)
2876.537 Million cell updates/aec

Title: US-09-674-277-19

Perfect score: 16

Sequence: 1 cggcatcgctgagtcg 16

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	93.8	275	9	LBAF040B07
2	15	93.8	323	9	AG241349
3	15	93.8	438	6	CB486369
4	15	93.8	456	6	CA995113
5	15	93.8	456	6	CA995466
6	15	93.8	458	6	CA996608
7	15	93.8	459	5	BQ125427
8	15	93.8	460	5	BQ626540
9	15	93.8	461	5	BQ626708
10	15	93.8	531	7	CF266988
11	15	93.8	562	9	TA43A03P
12	15	93.8	572	7	CV193301
13	15	93.8	595	1	AJ716704
14	15	93.8	765	8	BH438319
15	15	93.8	783	8	BH473894
16	14.4	90.0	91	1	AA486965
17	14.4	90.0	204	9	CG772409
18	14.4	90.0	301	4	BI306911
19	14.4	90.0	303	2	BE931552
20	14.4	90.0	326	6	CA548625
21	14.4	90.0	330	1	AI782868
22	14.4	90.0	376	9	FR002681.0
23	14.4	90.0	389	1	AA211806
24	14.4	90.0	401	1	AV403400

25	14.4	90.0	405	8	BZ302870
26	14.4	90.0	407	6	CB123076
27	14.4	90.0	410	7	R45538
28	14.4	90.0	414	8	BH254127
29	14.4	90.0	419	1	AA213737
30	14.4	90.0	422	7	CV163754
31	14.4	90.0	431	2	BF802567
32	14.4	90.0	432	2	BE772968
33	14.4	90.0	434	5	BW574947
34	14.4	90.0	456	4	BM270619
35	14.4	90.0	457	8	BZ516385
36	14.4	90.0	479	5	BX281216
37	14.4	90.0	480	1	AL039348
38	14.4	90.0	484	8	BZ303377
39	14.4	90.0	487	6	CD850585
40	14.4	90.0	491	6	CB176223
41	14.4	90.0	494	4	BJ158004
42	14.4	90.0	499	6	CD897040
43	14.4	90.0	500	6	CA341518
44	14.4	90.0	502	2	BE773039
45	14.4	90.0	504	6	CB276633

ALIGNMENTS

RESULT 1
LBAF040B07
LOCUS
DEFINITION
275 bp DNA linear GSS 20-JUN-2003
Leishmania braziliensis GSS, clone LBAF040B07, genomic survey
sequence.
ACCESSION
BX540045
VERSION
GI:32136956
KEYWORDS
GSS; genomic survey sequence.
SOURCE
Leishmania braziliensis
ORGANISM
Leishmania braziliensis
Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae;
Leishmania; Leishmania braziliensis species complex.
REFERENCE
1
Laurentino,E.C., Ruiz,J.C. and Cruz,A.K.
GSS analysis of the Leishmania braziliensis genome
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 275)
Cruz,A.K.
AUTHORS
Direct Submission
TITLE
Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo,
Department of Molecular and Cell Biology, FMRP, Avenida
Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL
JOURNAL
Clone requests: akcruz@mrp.usp.br.
COMMENT
Location/Qualifiers
FEATURES
source
1..275
/organism="Leishmania braziliensis"
/mol_type="genomic DNA"
/strain="MHOM/BR/75/M2904"
/db_xref="taxon:5660"
/clone="LBAF040B07"

ORIGIN
Query Match 93.8%; Score 15; DB 9; Length 275;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCATCGTCAGTTG 15

|||||
15 CGGCATCGTCAGTTG 29

RESULT 2

AG241349/c

LOCUS

DEFINITION

Lotus corniculatus var. japonicus DNA, clone:ljt08115_sfi, genomic survey sequence.

ACCESSION

AG241349

```

VERSION      AG241349.1  GI:26552186
KEYWORDS
SOURCE       Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
REFERENCE
1            Sato, S., Nakamura, Y. and Tabata, S.
AUTHORS      Lotus japonicus TAC End sequences
TITLE        Published Only in Database (2002)
JOURNAL
REFERENCE    2 (bases 1 to 323)
AUTHORS      Sato, S.
TITLE        Direct Submission
JOURNAL      Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
The First Laboratory for Plant Gene Research; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 252-0818, Japan
(E-mail: ssato@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/,
Tel: 81-438-52-3935 (ex. 2336), Fax: 81-438-52-3934)
LOCATION/Qualifiers
1. .323
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/strain="Miyakojima MG-20"
/variety="japonicus"
/db_xref="taxon:34305"
/clone="J1T08115_sfi"
/clone_lib="genomic TAC library"
/notes="VECTOR: pYLTAC7-synonym: Lotus japonicus"

ORIGIN
Query Match      93.8%; Score 15; DB 9; Length 323;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTG 15
    |||||
Db 300 CGGCATCGTCAGTTG 286

RESULT 3
LOCUS      CB486369
DEFINITION omykrtn001067 Oncorhynchus mykiss reproductive Oncorhynchus mykiss
          cDNA, mRNA sequence.
ACCESSION  CB486369
VERSION    CB486369.1  GI:29297595
KEYWORDS   EST.
SOURCE     Oncorhynchus mykiss (rainbow trout)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Euteleostei;
          Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE  1 (bases 1 to 438)
AUTHORS    GRASP Consortium, Davidson, W.S., Koop, B.F. and
          http://web.uvic.ca/cbr/grasp.
TITLE      A survey of Salmo salar transcripts from high complexity cDNA
          libraries
JOURNAL    Unpublished (2002)
COMMENT    Contact: Koop BF
          Centre for Biomedical Research
          University of Victoria
          PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
          Tel: 250 472 4067
          Fax: 250 472 4075
          Email: bkoop@uvic.ca
          Centre for Biomedical Research, University of Victoria cDNA
          preparation and sequencing: Roberto Alberto, Marianne
          Beet-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.
          bioinformatics: Gordon D Brown.
          Location/Qualifiers
          1. .438

FEATURES
source
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/strain="Karl Vierke"
/db_xref="taxon:8022"
/clone_lib="Oncorhynchus mykiss reproductive"
/notes="Vector: pcrtopo; Library Creator: Kristian R von
          Schauburg ; Rainbow trout tissue contributors: Mountain
          Trout Sales (Sooke, B.C.)"

ORIGIN
Query Match      93.8%; Score 15; DB 6; Length 438;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGCATCGTCAGTTGC 16
    |||||
Db 291 GGCATCGTCAGTTGC 305

RESULT 4
LOCUS      CA995113
DEFINITION t924c07.y1 Meloidogyne hapla J2 SL1 TOPO v1 Meloidogyne hapla cDNA
          5', mRNA sequence.
ACCESSION  CA995113
VERSION    CA995113.1  GI:27539984
KEYWORDS   EST.
SOURCE     Meloidogyne hapla
ORGANISM   Meloidogyne hapla
          Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
          Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE  1 (bases 1 to 456)
AUTHORS    McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
          Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
          Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
          Tsagarashvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
          Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
          Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
          McCann, R., Waterston, R. and Wilson, R.
          The Washington Univ. Nematode EST Project, 1999
          Unpublished (1999)
          Contact: McCarter JP
          The Washington Univ. Nematode EST Project, 1999
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          The library was constructed by Claire Murphy and Dr. James McCarter
          at Washington University, St. Louis. J2 were provided by Dr.
          Valerie Williamson of the University of California at Davis
          (vmwilliamson@ucdavis.edu).
          Putative full length read
          The vector to vector length is 457
          Seq primer: -40RP from Gibco.
          Location/Qualifiers
          1. .456
          /organism="Meloidogyne hapla"
          /mol_type="mRNA"
          /db_xref="taxon:6305"
          /dev_stage="J2"
          /lab_host="DH10B"
          /clone_lib="Meloidogyne hapla J2 SL1 TOPO v1"
          /notes="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;
          Site 2: EcoRI; The library was constructed by Claire
          Murphy and Dr. James McCarter at Washington University,
          St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne
          hapla J2 cDNA PCR products of size >400 nucleotides
          containing SL1 on the 5' end and oligo(dT) on the 3' end
          were non-directionally cloned into pCRII-TOPO(Invitrogen)
          following the TOPO TA cloning protocol. J2 were provided
          by Dr. Valerie Williamson of University of California at
          Davis (vmwilliamson@ucdavis.edu)."

FEATURES
source

```

ORIGIN

Query Match 93.8%; Score 15; DB 6; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16
 |||||

Db 286 GGCATCGTCAGTTGC 300

RESULT 5

CA995466

LOCUS

DEFINITION iq29f01.v1 Meloidogyne hapla J2 SL1 TOPO v1 Meloidogyne hapla cDNA

5, mRNA sequence.

CA995466

CA995466.1 GI:27540337

EST.

KEYWORDS

SOURCE

ORGANISM

Meloidogyne hapla

Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;

Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.

1 (bases 1 to 456)

McCarte, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,

Tagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,

Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

Contact: McCarte, J.P.

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter

at Washington University, St. Louis. J2 were provided by Dr.

Valerie Williamson of the University of California at Davis

(vmwilliamson@ucdavis.edu).

Putative full length read

The vector to vector length is 457

Seq primer: -40RP from Gibco.

Location/Qualifiers

1. .456

/organism="Meloidogyne hapla"

/mol_type="mRNA"

/db_xref="taxon:6305"

/dev_stage="J2"

/lab_host="DH10B"

/clone_lib="Meloidogyne hapla J2 SL1 TOPO v1"

/notes="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;

Site_2: EcoRI; The library was constructed by Claire

Murphy and Dr. James McCarter at Washington University,

St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne

hapla J2 cDNA PCR products of size >400 nucleotides

containing SL1 on the 5' end and oligo(dT) on the 3' end

were non-directionally cloned into pCRII-TOPO(Invitrogen)

following the ToPO TA cloning protocol. J2 were provided

by Dr. Valerie Williamson of University of California at

Davis (vmwilliamson@ucdavis.edu)."

ORIGIN

Query Match 93.8%; Score 15; DB 6; Length 456;

Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGCATCGTCAGTTGC 16

|||||

Db 286 GGCATCGTCAGTTGC 300

|||||

Db 288 GGCATCGTCAGTTGC 302

|||||

Db 288 GGCATCGTCAGTTGC 302

Db

286 GGCATCGTCAGTTGC 300

RESULT 6

CA996608

LOCUS

DEFINITION iq43e05.v1 Meloidogyne hapla J2 SL1 TOPO v1 Meloidogyne hapla cDNA

5, mRNA sequence.

CA996608

CA996608.1 GI:27541479

EST.

KEYWORDS

SOURCE

ORGANISM

Meloidogyne hapla

Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;

Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.

1 (bases 1 to 458)

McCarte, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,

Tagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,

Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

Contact: McCarte, J.P.

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter

at Washington University, St. Louis. J2 were provided by Dr.

Valerie Williamson of the University of California at Davis

(vmwilliamson@ucdavis.edu).

Putative full length read

The vector to vector length is 459

Seq primer: -40RP from Gibco.

Location/Qualifiers

1. .458

/organism="Meloidogyne hapla"

/mol_type="mRNA"

/db_xref="taxon:6305"

/dev_stage="J2"

/lab_host="DH10B"

/clone_lib="Meloidogyne hapla J2 SL1 TOPO v1"

/notes="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;

Site_2: EcoRI; The library was constructed by Claire

Murphy and Dr. James McCarter at Washington University,

St. Louis. Oligo(dT)-SL1 PCR based library. Meloidogyne

hapla J2 cDNA PCR products of size >400 nucleotides

containing SL1 on the 5' end and oligo(dT) on the 3' end

were non-directionally cloned into pCRII-TOPO(Invitrogen)

following the ToPO TA cloning protocol. J2 were provided

by Dr. Valerie Williamson of University of California at

Davis (vmwilliamson@ucdavis.edu)."

RESULT 7

BQ125427

LOCUS

DEFINITION rc68b06.v1 Meloidogyne hapla egg SL1 TOPO v1 Meloidogyne hapla

cDNA, mRNA sequence.

QY 2 GGCATCGTCAGTTGC 16

|||||

Db 288 GGCATCGTCAGTTGC 302

|||||

```

ACCESSION      B0125427
VERSION        B0125427.1  GI:20199338
KEYWORDS       EST.
SOURCE         Meloiodogyne hapla
ORGANISM       Meloiodogyne hapla

REFERENCE
AUTHORS        McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
               Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
               Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
               Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
               Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
               Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
               McCann,R., Waterston,R. and Wilson,R.
               The Washington Univ. Nematode EST Project, 1999
               Unpublished (1999)

TITLE          The Washington Univ. Nematode EST Project, 1999
JOURNAL
COMMENT        The Washington Univ. Nematode EST Project, 1999
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: est@watson.wustl.edu
               The library was constructed by Claire Murphy and Dr. James McCarter
               at Washington University, St. Louis. Eggs were provided by Dr.
               David Bird of North Carolina State University, Raleigh, NC
               (david.bird@ncsu.edu). DNA Sequencing by: Washington University
               Genome Sequencing Center

FEATURES
source         Putative full length read
               The vector to vector length is 460
               Seq primer: SL1 primer.
               Location/Qualifiers
                 1..459
                 /organism="Meloiodogyne hapla"
                 /mol_type="mRNA"
                 /db_xref="taxon:6305"
                 /dev_stage="eggs"
                 /lab_host="DH10B"
                 /clone_lib="Meloiodogyne hapla egg SL1 TOPO v1"
                 /note="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;
                 Site 2: EcoRI; The library was constructed by Claire
                 Murphy and Dr. James McCarter at Washington University,
                 St. Louis. Oligo(dT)-SL1 PCR based library. Meloiodogyne
                 hapla egg cDNA PCR products of size >400 nucleotides
                 containing SL1 on the 5' end and oligo(dT) on the 3' end
                 were non-directionally cloned into pCRII-TOPO(Invitrogen)
                 following the Topo TA cloning protocol. Eggs were provided
                 by Dr. David Bird of North Carolina State University,
                 Raleigh, NC (david.bird@ncsu.edu)"

ORIGIN
Query Match    93.8%; Score 15; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2  GGCATCGTCAGTTGC 16
      |||||
Db   289 GGCATCGTCAGTTGC 303

RESULT 8
BQ626540
LOCUS      BQ626540
DEFINITION p21le03.v1 Pratylenchus penetrans mixed stage SL1 TOPO v1
            Pratylenchus penetrans cDNA 5', mRNA sequence.
ACCESSION  BQ626540
VERSION    BQ626540.1  GI:21653718
KEYWORDS   EST.
SOURCE     Pratylenchus penetrans
ORGANISM   Pratylenchus penetrans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;

Query Match    93.8%; Score 15; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2  GGCATCGTCAGTTGC 16
      |||||
Db   289 GGCATCGTCAGTTGC 303

RESULT 9
BQ626708
LOCUS      BQ626708
DEFINITION p213R07.y1 Pratylenchus penetrans mixed stage SL1 TOPO v1
            Pratylenchus penetrans cDNA 5', mRNA sequence.
ACCESSION  BQ626708
VERSION    BQ626708.1  GI:21653886
KEYWORDS   EST.
SOURCE     Pratylenchus penetrans
ORGANISM   Pratylenchus penetrans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Pratylenchidae; Pratylenchinae; Pratylenchus.
REFERENCE
AUTHORS        McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
               Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
               Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
               Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
               Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
               Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
               McCann,R., Waterston,R. and Wilson,R.
               The Washington Univ. Nematode EST Project, 1999
               Unpublished (1999)

TITLE          The Washington Univ. Nematode EST Project, 1999
JOURNAL
COMMENT        The Washington Univ. Nematode EST Project, 1999
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: est@watson.wustl.edu
               The library was constructed by Claire Murphy and Dr. James McCarter
               at Washington University, St. Louis. RNA was provided by Andrew
               Kloeck of Divergence Inc., St. Louis, MO.
               Putative full length read
               The vector to vector length is 461
               Seq primer: SL1 primer.
               Location/Qualifiers
                 1..460
                 /organism="Pratylenchus penetrans"
                 /mol_type="mRNA"
                 /db_xref="taxon:45929"
                 /dev_stage="mixed stage"
                 /lab_host="DH10B (Invitrogen)"
                 /clone_lib="Pratylenchus penetrans mixed stage SL1 TOPO
                 v1"
                 /note="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;
                 Site 2: EcoRI; The library was constructed by Claire
                 Murphy and Dr. James McCarter at Washington University,
                 St. Louis. Oligo(dT)-SL1 PCR based library. cDNA PCR
                 products of size >400 nucleotides containing SL1 on the 5'
                 end and oligo(dT) on the 3' end were non-directionally
                 cloned into pCRII-TOPO(Invitrogen) following the Topo TA
                 cloning protocol. RNA was provided by Andrew Kloeck of
                 Divergence, Inc., St. Louis, MO."

```


TITLE
JOURNAL
COMMENT

The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. RNA was provided by Andrew Kloek of Divergence Inc., St. Louis, MO.
Putative full length read
The vector to vector length is 462

Seq primer: SL1 primer
Location/Qualifiers
1. .461
/organism="Pratylenchus penetrans"
/mol_type="mRNA"
/db_xref="taxon:45929"
/dev_stage="mixed stage"
/lab_host="DH10B (Invitrogen)"
/clone_lib="Pratylenchus penetrans mixed stage SL1 TOPO v1"

FEATURES
source

/note="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI; Site 2: EcoRI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the TOPO TA cloning protocol. RNA was provided by Andrew Kloek of Divergence, Inc., St. Louis, MO."

ORIGIN

Query Match 93.8%; Score 15; DB 5; Length 461;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GGCATCGTCAGTTGC 16
|||||||
Db 291 GGCATCGTCAGTTGC 305

RESULT 10
CF266988/c
LOCUS
DEFINITION
CF266988
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
COMMENT

CF266988 531 bp mRNA linear EST 13-AUG-2003
cDNA clone TgESTzyj21a08.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii
CF266988
CF266988.1 GI:33628941
EST.
Toxoplasma gondii
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
1 (bases 1 to 531)
Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M., Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.
Toxoplasma EST Project
Unpublished (2001)
Contact: Clifton, S.
Toxoplasma EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@wustl.edu
Contact David Sibley (toxost@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.

TITLE
JOURNAL
COMMENT

The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. RNA was provided by Andrew Kloek of Divergence Inc., St. Louis, MO.
Putative full length read
The vector to vector length is 462

Seq primer: SL1 primer
Location/Qualifiers
1. .531
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="Tachyzoite"
/db_xref="taxon:5811"
/clone_lib="TgESTzyj21a08.y1"
/dev_stage="Tachyzoite"
/lab_host="Electrofen Blue cells (Stratagene)"
/clone_lib="TgMAS Tachyzoite cDNA Library"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Keliang Tang, and Robert Cole at Washington University. cDNA was synthesized from Poly(A)+ mRNA using an oligo-d(T) primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on Sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into Electrofen Blue cells (Stratagene). The library may contain a small percentage of host or bacterial contaminants."

ORIGIN

Query Match 93.8%; Score 15; DB 7; Length 531;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGGCATCGTCAGTTG 15
|||||||
Db 41 CGGCATCGTCAGTTG 27

RESULT 11
TA43A03P/c
LOCUS
DEFINITION
TA43A03P
T. brucei sheared genomic DNA clone 43a03, forward sequence, genomic survey sequence.
AL454668
AL454668.1 GI:11856292
GSS.
Trypanosoma brucei
Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
1 (bases 1 to 562)
Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre. The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 Gutat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
Location/Qualifiers
1. .562
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"

Db	125 CGGCATCGTCAGTTG 111		Db	1 CGGCATCGTCAGTTG 15	
RESULT 13	AJ716704			RESULT 12	CV193301/c
LOCUS	AJ716704			LOCUS	CV193301
DEFINITION	Triticum turgidum subsp. durum cDNA clone 05257R, mRNA sequence.			DEFINITION	SnESTab28h01.Y1 Sarcocystis neurona merozoite UK cSN4 1 cDNA library Sarcocystis neurona cDNA 5', mRNA sequence.
ACCESSION	AJ716704			ACCESSION	CV193301
VERSION	AJ716704.1 GI:49603286			VERSION	CV193301.1 GI:52122138
KEYWORDS	EST.			KEYWORDS	EST.
SOURCE	Triticum turgidum subsp. durum (durum wheat)			SOURCE	Sarcocystis neurona
ORGANISM	Triticum turgidum subsp. durum			ORGANISM	Sarcocystis neurona
REFERENCE	1 (bases 1 to 595)			REFERENCE	1 (bases 1 to 572)
AUTHORS	Cifarelli, R.A., D'Onofrio, O. and Lauria, G.			AUTHORS	Howe, D.K., Stamper, S., Tang, K., Sibley, L.D., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wyllie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., McCann, R., Blistain, A., Bennet, J., Schmitt, A., Ronko, I., Tsagareishvili, R., Fedele, M., Belaygorod, L., Franklin, C., Carr, L.M., Grow, A., Maguire, L., Wadkins, J., Richey, J., Waterston, R. and Wilson, R.
TITLE	Expressed Sequences Tags (ESTs) library from totipotent cDNA of durum wheat			TITLE	Sarcocystis neurona EST project
JOURNAL	Unpublished (2003)			JOURNAL	Unpublished (2000)
COMMENT	Contact: Cifarelli RA			COMMENT	Contact: Daniel K. Howe
FEATURES	Location/Qualifiers			FEATURES	Location/Qualifiers
source	1..595			source	1..572
ORIGIN	/organism="Triticum turgidum subsp. durum"			ORIGIN	/organism="Sarcocystis neurona"
Query Match	93.8%; Score 15; DB 1; Length 595;			Query Match	93.8%; Score 15; DB 7; Length 572;
Best Local Similarity	100.0%; Pred. No. 1.1e+03;			Best Local Similarity	100.0%; Pred. No. 1.1e+03;
Matches	15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			Matches	15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CGGCATCGTCAGTTG 15			Qy	1 CGGCATCGTCAGTTG 15
Db	413 CGGCATCGTCAGTTG 427			Db	1 CGGCATCGTCAGTTG 15
RESULT 14	BH438319			RESULT 14	BH438319
LOCUS	BH438319			LOCUS	BH438319
DEFINITION	BOGH102TR BOGH Brassica oleracea genomic clone BOGH102, genomic survey sequence.			DEFINITION	BOGH102TR BOGH Brassica oleracea genomic clone BOGH102, genomic survey sequence.
ACCESSION	BH438319			ACCESSION	BH438319
VERSION	BH438319.1 GI:17624040			VERSION	BH438319.1 GI:17624040
KEYWORDS	GSS.			KEYWORDS	GSS.
SOURCE	Brassica oleracea			SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea			ORGANISM	Brassica oleracea
REFERENCE	1 (bases 1 to 765)			REFERENCE	1 (bases 1 to 765)
AUTHORS	Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.			AUTHORS	Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE	Whole genome shotgun sequencing of Brassica oleracea			TITLE	Whole genome shotgun sequencing of Brassica oleracea
JOURNAL	Unpublished (2001)			JOURNAL	Unpublished (2001)
COMMENT	Contact: Chris Town			COMMENT	Contact: Chris Town
Qy	1 CGGCATCGTCAGTTG 15			Qy	1 CGGCATCGTCAGTTG 15

```

FEATURES
  source
    Class: sheared ends.
    Location/Qualifiers
      1..765
        /organism="Brassica oleracea"
        /mol_type="genomic DNA"
        /strain="TO1000DH3"
        /db_xref="taxon:3712"
        /clone="BOGHI02"
        /clone_lib="BOGH"
        /note="vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
        genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN
  Query Match      93.8%; Score 15; DB 8; Length 765;
  Best Local Similarity 100.0%; Pred. No. 1.1e+03;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTG 15
    |||||
Db 314 CGGCATCGTCAGTTG 328

RESULT 15
BH473894
LOCUS
DEFINITION
  BOGUB33TR BOGU Brassica oleracea genomic clone BOGUB33, genomic
  survey sequence.
ACCESSION
  BH473894
VERSION
  BH473894.1 GI:17682005
KEYWORDS
  GSS.
SOURCE
  Brassica oleracea
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 783)
  Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Fraser, C.M.
  Whole genome shotgun sequencing of Brassica oleracea
  Unpublished (2001)
  Other_GSSs: BOGUB33TF
  Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Tel: 301-838-3523
  Fax: 301-838-0208
  Email: cdtown@tigr.org
  DNA is from a doubled haploid provided by Tom Osborn.
  Seq primer: TR
  Class: sheared ends.
  Location/Qualifiers
    1..783
      /organism="Brassica oleracea"
      /mol_type="genomic DNA"
      /strain="TO1000DH3"
      /db_xref="taxon:3712"
      /clone="BOGUB33"
      /clone_lib="BOGU"
      /note="vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
      genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN
  Query Match      93.8%; Score 15; DB 8; Length 783;
  Best Local Similarity 100.0%; Pred. No. 1.1e+03;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTG 15
    |||||
Db 148 CGGCATCGTCAGTTG 162

RESULT 16
AA486965/c
LOCUS
  AA486965
  91 bp mRNA linear EST 06-MAR-1998

```

```

DEFINITION
  ab17f07 ri Stratagene lung (#937210) Homo sapiens cDNA clone
  IMAGE:841093 5' similar to TR:G1146334 G1146334 LIN19 PROTEIN. ;,
  mRNA sequence.
ACCESSION
  AA486965
VERSION
  AA486965.1 GI:2217129
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 91)
  Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
  Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
  Martin, J., Moore, B., Schellenger, K., Steptoe, M., Tan, F.,
  Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
  WashU-NCI human EST Project
  Unpublished (1997)
  Contact: Wilson RK
  Washington University School of Medicine
  444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available royalty-free through LNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Trace considered overall poor quality
  Possible reversed clone: similarity on wrong strand
  Insert Length: 1446 Std Error: 0.00
  Seq primer: -28m13 rev1 ET from Amersham
  High quality sequence stop: 1.
  Location/Qualifiers
    1..91
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:841093"
      /sex="male"
      /dev_stage="72 years"
      /lab_host="SOLR cells (kanamycin resistant)"
      /clone_lib="Stratagene lung (#937210)"
      /note="Organ: lung; Vector: pBluescript SK-; Site 1:
      EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
      Oligo dT. normal lung. Average insert size: 1.0 kb;
      Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG
      3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
  Query Match      90.0%; Score 14.4; DB 1; Length 91;
  Best Local Similarity 93.8%; Pred. No. 1.9e+03;
  Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
    |||||
Db 35 CGGCATCGTCAGTTGC 20

RESULT 17
CG772409/c
LOCUS
  CG772409
  204 bp DNA linear GSS 29-OCT-2003
  DEFINITION
    1123009G01.yl 1123 - RescueMu Grid L Zea mays genomic, genomic
    survey sequence.
  ACCESSION
    CG772409
  VERSION
    CG772409.1 GI:38026360
  KEYWORDS
    GSS.
  SOURCE
    Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 204)
  Walbot, V.
  Maize genomic sequences found using engineered RescueMu transposon
  Unpublished (2001)

```

Mon Mar 14 11:04:18 2005

```

COMMENT
Contact: Walbot V
Department of Biological Sciences
Stanford University
955 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site so sequence was trimmed. Post-ligation
sequence submitted separately.
Plate: 1123009 row: 10
Class: transposon-tagged.
Location/Qualifiers
1. .204
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73/K55"
/db_xref="taxon:4577"
/dev_stage="adult"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_lib="1123 - RescueMu Grid L"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
RescueMu is a 4.9 Kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu'. Grid L was grown in Molokai in 2001. DNA was
extracted from leaf strips, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."

ORIGIN
Query Match 90.0%; Score 14.4; DB 9; Length 204;
Best Local Similarity 93.8%; Pred. No. 2.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
|||||
Db 41 CGGCATCGGCAGTTGC 26

RESULT 18
BI306911 301 bp mRNA linear EST 31-DEC-2001
LOCUS PT50194 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
tricornutum cDNA 5', mRNA sequence.
ACCESSION BI306911
VERSION BI306911.1 GI:18020615
KEYWORDS EST.
SOURCE Phaeodactylum tricornutum
ORGANISM Phaeodactylum tricornutum
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 301)
Scala,S., Carls,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
Genome properties of the diatom Phaeodactylum tricornutum
Plant Physiol. 129 (3), 993-1002 (2002)
2211123
PUBMED 12114555
COMMENT Contact: Bowler C
Laboratory of Molecular Plant Biology
Stazione Zoologica 'Anton Dohrn'
Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: chris@alpha.szn.it
Seq primer: T3 backward.
Location/Qualifiers
1. .301
/organism="Phaeodactylum tricornutum"

FEATURES
source
Query Match 90.0%; Score 14.4; DB 9; Length 204;
Best Local Similarity 93.8%; Pred. No. 2.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
|||||
Db 41 CGGCATCGGCAGTTGC 26

RESULT 19
BE931552/c 303 bp mRNA linear EST 02-OCT-2000
LOCUS RC4-HT0444-250800-012-d12 HT0444 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE931552
ACCESSION BE931552
VERSION BE931552.1 GI:10457628
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 303)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.P., Mateukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=8t2=RC4-HT0444-250
800-012-d12&t3=2000-08-25&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 47
High quality sequence stop: 303.
Location/Qualifiers
1. .303
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0444"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 90.0%; Score 14.4; DB 2; Length 303;
Best Local Similarity 93.8%; Pred. No. 2.2e+03;

```

```

/mol_type="mRNA"
/db_xref="taxon:2850"
/cell_line="CCMP632"
/clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
/note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
Xho I"

```

ORIGIN

```

Query Match 90.0%; Score 14.4; DB 4; Length 301;
Best Local Similarity 93.8%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 CGGCATCGTCAGTTGC 16
|||||

```

```

Db 168 CGGCATCTTCAGTTGC 183
|||||

```

RESULT 19

```

BE931552/c 303 bp mRNA linear EST 02-OCT-2000
LOCUS RC4-HT0444-250800-012-d12 HT0444 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE931552
ACCESSION BE931552
VERSION BE931552.1 GI:10457628
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 303)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.P., Mateukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=8t2=RC4-HT0444-250
800-012-d12&t3=2000-08-25&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 47
High quality sequence stop: 303.
Location/Qualifiers
1. .303
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0444"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

FEATURES

source

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 303 CGGCATCGTCAGTTGC 288

RESULT 20
 CA548625/c
 LOCUS
 DEFINITION
 C0808A11-5N NIA Mouse Blastocyst cDNA Library (Long) EST 19-NOV-2002
 cDNA clone NIA:C0808A11 IMAGE:30025738 5', mRNA sequence.
 CA548625
 VERSION
 CA548625.1 GI:35092522
 EST.

SOURCE
 Mus musculus (house mouse)

ORGANISM

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS
 1 (bases 1 to 326)
 Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Tanaka, T., Luo, A. and
 Ko, M.S.H.

TITLE
 Systematic Analyses of NIA Mouse Blastocyst cDNA Library (Long)

JOURNAL

COMMENT
 Laboratory of Genetics

National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov

Plate: C0808 row: A column: 11

Seq primer: M13 Reverse

High quality sequence stop: 326

POLYA=No.

FEATURES

source
 Location/Qualifiers

1..326

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="niabst:C0808A11-5N"

/db_xref="taxon:10090"

/clone="NIA:C0808A11 IMAGE:30025738"

/tissue_type="Blastocyst"

/dev_stage="3.5-dpc"

/lab_host="DH10B"

/clone_lib="NIA Mouse Blastocyst cDNA Library (Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is

a long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

extracted from a pool of 20 Blastocysts. Double-stranded

cDNAs were synthesized with an Oligo(dT) primer

[Invitrogen]:

5'-PGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3' from

0.2 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated

to lona-linker L1-Sal4, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Taq polymerase (Takara) with a primer Sal4-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes

and cloned into SalI/NotI site of pSPORT1 plasmid vector.

The DH10B E. coli host was transformed with the ligation

mixture by the standard chemical method. The average

insert size is about 2.2 kb. The library was constructed

by Yulan Piao (NIA)."

ORIGIN

Query Match 90.0%; Score 14.4; DB 6; Length 326;
 Best Local Similarity 93.8%; Pred. No. 2.2e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16
 |||||
 Db 29 CGGCATCGTCAGTTGC 14

RESULT 21

AI782868

LOCUS

DEFINITION

61400SH02.x2 614 - root cDNA library from Walbot Lab Zea mays cDNA,

mRNA sequence.

ACCESSION

AI782868

VERSION

AI782868.1 GI:5296188

KEYWORDS

EST.

SOURCE

ORGANISM

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 330)

Walbot, V.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 614005 row: H column: 02.

Location/Qualifiers

1..330

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="W23"

/db_xref="taxon:4577"

/tissue_type="root"

/dev_stage="3-4 days old"

/lab_host="XL0LR"

/clone_lib="614 - root cDNA library from Walbot Lab"

/note="Organ: root; Vector: pBluescriptII SK+; Site 1:

EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot

Lab (LM)"

ORIGIN

Query Match 90.0%; Score 14.4; DB 1; Length 330;

Best Local Similarity 93.8%; Pred. No. 2.2e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

|||||

Db 251 CGGCATCGTCAGTTGC 266

RESULT 22

FR0026810

LOCUS

DEFINITION

F.rubripes GSS sequence, clone 180N03AC11, genomic survey sequence.

ACCESSION

AL019639

VERSION

AL019639.1 GI:2686007

KEYWORDS

GSS; genome survey sequence.

SOURCE

Takifugu rubripes (Fugu rubripes)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;

Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE

AUTHORS

1

Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,

Bouchred, N., Cottage, A., Yeo, G.S., Umrana, Y., Williams, G. and

Brenner, S.

Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
Genome Res. 9 (10), 960-971 (1999)

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

2 (bases 1 to 376)
Elgar,G., Clark M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.

TITLE
JOURNAL

Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmpr.mrc.ac.uk

COMMENT

Vector: pBluescript II KS

V type: phagemid

PRIMER: KS

DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic

sequence.

FEATURES

source

Location/Qualifiers
1..376
/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="180N03aC11"
/clone_lib="cosmid 180N03"

ORIGIN

Query Match 90.0%; Score 14.4; DB 9; Length 376;
Best Local Similarity 93.8%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 56 CGGCATCGACAGTTGC 71

RESULT 23

AA211806/c

LOCUS

zr99c07.r1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:682860 5'
similar to TR:G1146334 G1146334 LIN19 PROTEIN. ;, mRNA sequence.

ACCESSION

AA211806

VERSION

AA211806.1 GI:1810442

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 389)

AUTHORS

NCI-CGAP

TITLE

NCI-CGAP

http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m3 rev2 ET from Amersham

High quality sequence stop: 299.

Location/Qualifiers

1..389

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:5586180"

/db_xref="taxon:9606"

/clone="IMAGE:682860"

/tissue_type="germinal center B cell"

/lab_host="DH108"

/clone_lib="NCI CGAP GCBI"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTCAAGTGGGCGGCTCATTTTTTTTTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 90.0%; Score 14.4; DB 1; Length 389;

Best Local Similarity 93.8%; Pred. No. 2.3e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 47 CGGCATCGTCACATTGC 32

RESULT 24

AV403400

LOCUS

DEFINITION

AV403400 Bombyx mori middle silkland 5th-instar larva Bombyx mori

cDNA clone msgV0424 T3, mRNA sequence.

ACCESSION

AV403400

VERSION

AV403400.1 GI:6907488

KEYWORDS

EST.

SOURCE

Bombyx mori (domestic silkworm)

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;

Bombycoidea; Bombycidae; Bombyx.

REFERENCE

1 (bases 1 to 401)

AUTHORS

Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.

Bombyx mori cDNA

Unpublished (2000)

CONTACT: Mita K

Genome Research Group

National Institute of Agrobiological Sciences

Onsai 1-2, Tsukuba, Ibaraki 305-8634, Japan

Email: kmita@nias.affrc.go.jp

method: uni-directional, sequence direction: sequenced from T3 primer

(5' -> 3')

Project="Silkworm Genome Program in MAFF, and Research for the

Future Program in JSPS". see 'SilkBase',

<http://www.silkbase.ac.jp/silkbase/>, for whole ESTdb.

Location/Qualifiers

1..401

/organism="Bombyx mori"

/mol_type="mRNA"

/db_xref="taxon:7091"

/clone="msgV0424"

/sex="female/male mixed"

/tissue_type="middle silkland"

/dev_stage="5th-instar larva"

/clone_lib="Bombyx mori middle silkland 5th-instar larva"

/note="Donated by Dr. Sehnal, Czech"

ORIGIN

Query Match 90.0%; Score 14.4; DB 1; Length 401;

Best Local Similarity 93.8%; Pred. No. 2.3e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGCATCGTCAGTTGC 16

Db 19 CGGCATCATCATGTTGC 34

RESULT 25

BZ302870

LOCUS

DEFINITION

BZ302870 Kluyveromyces delphensis Random Genomic Library

KD2075.q1 Kluyveromyces delphensis genomic clone KD2075, genomic survey

provided by Dr. David Allman

sequence.
ACCESSION B2302870
VERSION B2302870.1 GI:24450339
KEYWORDS GSS.
SOURCE Kluyveromyces delphensis
ORGANISM Kluyveromyces delphensis
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
TITLE 1 (bases 1 to 405)
JOURNAL Wong, S., Fares, M.A., Zimmermann, W., Butler, G. and Wolfe, K.H.
MEDLINE Evidence from comparative genomics for a complete sexual cycle in
PUBMED the 'asexual' pathogenic yeast Candida glabrata
COMMENT Genome Biol. 4 (2), R10 (2003)
22508158
12620120
Contact: Wong S
Department of Genetics, Smurfit Institute
Trinity College Dublin
Dublin 2, Ireland
Tel: 353 1 6082119
Fax: 353 1 6798558
Email: swong@tcd.ie
Class: plasmid ends.
FEATURES
source
1..405
Location/Qualifiers
/organism="Kluyveromyces delphensis"
/mol_type="genomic DNA"
/strain="CBS 2170"
/db_xref="taxon:51657"
/clone="KD2075"
/clone_lib="Kluyveromyces delphensis Random Genomic Library"
ORIGIN
Query Match 90.0%; Score 14.4; DB 8; Length 405;
Best Local Similarity 93.8%; Pred.No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 CGGCATCCTCAGTTGC 16
|||||
Db 120 CGGCATCTTCAGTTGC 135
|||||

Search completed: March 11, 2005, 13:01:17
Job time : 213.723 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 02:02:00 ; Search time 26.3327 Seconds
(without alignments)
4046.498 Million cell updates/sec

Title: US-09-674-277-20

Perfect score: 18

Sequence: 1 acgcgcgtcagttgcs 18

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04.*
1: Geneseq_1980s.*
2: Geneseq_1990s.*
3: Geneseq_2000s.*
4: Geneseq_2001s.*
5: Geneseq_2001bs.*
6: Geneseq_2002s.*
7: Geneseq_2002bs.*
8: Geneseq_2003s.*
9: Geneseq_2003bs.*
10: Geneseq_2003cs.*
11: Geneseq_2003ds.*
12: Geneseq_2004s.*
13: Geneseq_2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	3	AZ36120 Primer de
2	18	100.0	31	3	AZ36112 Primer de
3	18	100.0	31	3	AZ36113 Primer de
4	18	100.0	1489	3	AZ36101 Nucleic a
C 5	16.4	91.1	664	10	ADB68745 Pseudomon
C 6	16.4	91.1	1056	13	ADS46729 Bacterial
7	16.4	91.1	1137	6	ABZ13613 Arabidops
8	16.4	91.1	1301	3	AC33636 Arabidops
9	16.4	91.1	1362	13	ADT46111 Bacterial
10	16	88.9	16	3	AZ36119 Primer de
11	15.4	85.6	930	8	ACA40338 Prokaryot
12	15.4	85.6	936	8	ACA38540 Prokaryot
C 13	15.4	85.6	1392	8	ACA36906 Prokaryot
C 14	15.4	85.6	1563	10	ACF70905 Phototrab
C 15	15.4	85.6	1632	2	AAQ96251 Fructosyl
C 16	15.4	85.6	1797	10	AAZ57635 Rice dise
17	15.4	85.6	22934	4	AAZ59613 Propionib
18	15.4	85.6	22934	8	ACF64542 Propionib
C 19	15.4	85.6	25860	11	ADR20887 Phototrab
20	15.4	85.6	110000	4	AAI99682_05 Continuation (6 of

21	15.4	85.6	110000	4	AAI99682_06 Continuation (7 of
22	15.4	85.6	110000	4	AAI99683_05 Continuation (6 of
23	15.4	85.6	110000	4	AAI99683_06 Continuation (7 of
C 24	15.4	85.6	110000	6	ABA03041_06 Continuation (41 o
C 25	15.4	85.6	110000	10	ACF67367_40 Continuation (9 of
26	15.4	85.6	110000	10	ACF65388_08 Adk5969 Plant DNA
27	15	83.3	319	10	ADK56969 Plant DNA
28	15	83.3	319	10	ADK55125 Plant DNA
29	15	83.3	340	10	ADK53158 Plant DNA
30	15	83.3	368	10	ADK56429 Plant DNA
C 31	14.8	82.2	35	8	ACA10036 Necrosis
C 32	14.8	82.2	38	6	ACN28184 WNV minus
C 33	14.8	82.2	38	6	ACN30046 WNV minus
C 34	14.8	82.2	38	6	ACN16388 WNV Hamme
C 35	14.8	82.2	38	6	ACN28717 WNV minus
C 36	14.8	82.2	38	6	ACN27201 WNV minus
C 37	14.8	82.2	38	8	ACA07388 Necrosis
C 38	14.8	82.2	38	8	ACD51549 HBV hamme
C 39	14.8	82.2	38	11	ADL56456 Human PKR
C 40	14.8	82.2	38	11	ADL75523 Human PTG
C 41	14.8	82.2	38	12	ADM61015 Hepatitis
42	14.8	82.2	338	6	AAZ20030 Human CDN
C 43	14.8	82.2	346	12	ACH91414 Human gen
44	14.8	82.2	381	3	AAZ82016 N. mening
45	14.8	82.2	420	8	ACA41396 Prokaryot

ALIGNMENTS

RESULT 1
AAZ36120
ID AAZ36120 standard; DNA; 18 BP.
XX
AC AAZ36120;
XX
DT 11-FEB-2000 (first entry)
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; katp gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
OS Synthetic.
OS Escherichia coli.
XX
PN WO9955908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
FA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
DR WPI; 2000-013443/01.
XX
PT New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 5; Page 27; 48pp; French.
XX
CC AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katp gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 18 BP; 3 A; 5 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 3; Length 18;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
 |||||
 Db 1 ACGGCATCGTCAGTTGCG 18

RESULT 2
 AAZ36112
 ID AAZ36112 standard; DNA; 31 BP.

XX AC AAZ36112;

XX DT 11-FEB-2000 (first entry)

XX DE Primer derived from a nucleic acid sequence specific to EHEC.

XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 XX PCR primer; probe; ss.

XX OS Synthetic.
 XX OS Escherichia coli.

XX FN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX PS WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
 XX Escherichia coli, particularly serotype O157:H7, used for detecting these
 XX bacteria in food.

XX PS Claim 5; Page 27; 48pp; French.

XX CC AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 3; Length 31;

Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
 |||||
 Db 5 ACGGCATCGTCAGTTGCG 22

RESULT 3

AAZ36113

ID AAZ36113 standard; DNA; 31 BP.

XX AC AAZ36113;

XX DT 11-FEB-2000 (first entry)

XX DE Primer derived from a nucleic acid sequence specific to EHEC.

XX KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 XX IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 XX PCR primer; probe; ss.

XX OS Synthetic.
 XX OS Escherichia coli.

XX FN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX PS WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
 XX Escherichia coli, particularly serotype O157:H7, used for detecting these
 XX bacteria in food.

XX PS Claim 5; Page 27; 48pp; French.

XX CC AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 18; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
 |||||
 Db 9 ACGGCATCGTCAGTTGCG 26

RESULT 4

AAZ36101

ID AAZ36101 standard; DNA; 1489 BP.

```

XX - AA236101;
XX
XX
DT 11-FEB-2000 (first entry)
XX
XX
DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.
XX
XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; ds.
XX
XX Escherichia coli.
OS
XX
XX WO955908-A2.
PN
XX
XX 04-NOV-1999.
PD
XX
XX 27-APR-1999; 99WO-FR001000.
PF
XX
XX 28-APR-1998; 98FR-00005329.
XX
XX (SNFI ) PASTEUR SANOFI DIAGNOSTICS.
XX
XX Frechon DTM, Laure FC, Thierry D;
PI
XX
XX WPI; 2000-013443/01.
DR
XX
XX New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
XX Claim 1; Fig 1; 48pp; French.
XX
XX The present sequence is specific to enterohemorrhagic Escherichia coli
CC (EHEC). The sequence is 99.9% homologous to the katP gene of E. coli
CC O157:H7 (nucleotides 407-1489 of the present sequence), and 95.8%
CC homologous with IS91 of E. coli (nucleotides 1-406 of the present
CC sequence). The present sequence is of plasmid origin. Fragments of the
CC present sequence are used, as probes and primers, for detection of E.
CC coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
CC animal samples, foods or the environment. The fragments are also useful
CC for epidemiological studies
XX
XX Sequence 1489 BP; 386 A; 354 C; 391 G; 358 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 18; DB 3; Length 1489;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACGGCATCGTCAGTTGCG 18
DB ||||||||||||||||
395 ACGGCATCGTCAGTTGCG 412
RESULT 5
ADB68745/c
ID ADB68745 standard; DNA; 664 BP.
XX
XX ADB68745;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Pseudomonas chlororaphis phzI DNA.
DE
XX
XX quorum sensing; lux homologue; luxI; ds.
KW
XX
XX Pseudomonas chlororaphis.
OS
XX
XX WO2003057902-A2.
PN
XX
XX 17-JUL-2003.
PD
XX
XX 08-JAN-2003; 2003WO-US000479.
PF
XX
XX

```

```

PR 08-JAN-2002; 2002US-0346531P.
PR 07-JAN-2003; 2003US-00338110.
XX
XX (FRAU ) FRAUNHOFER USA INC.
PA
XX
XX Fuhrmann JJ, Romeiser JA;
PI
XX
XX WPI; 2003-618102/58.
DR
XX
XX Detecting quorum sensing potential of a Gram-negative bacterium in a
PT sample comprises performing a polymerase chain reaction using nucleic
PT acids extracted from a sample containing a microorganism.
XX
XX Disclosure; Fig 10; 86pp; English.
PS
XX
XX The invention relates to a novel method for detecting the quorum sensing
CC potential of a microorganism in a sample which comprises performing PCR
CC using nucleic acids extracted from a sample containing at least one type
CC of microorganism. The method may be useful for detecting the quorum
CC sensing potential of a microorganism in a sample by amplifying a fragment
CC of a lux gene or homologue. The current sequence is that of the luxI
CC homologue DNA of the invention.
XX
XX Sequence 664 BP; 162 A; 219 C; 167 G; 116 T; 0 U; 0 Other;
SQ
Query Match 91.1%; Score 16.4; DB 10; Length 664;
Best Local Similarity 94.4%; Pred. No. 95;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACGGCATCGTCAGTTGCG 18
DB ||||||||||||||||
419 ACGGCATCGTCAGTTGCG 402
RESULT 6
ADS46729/c
ID ADS46729 standard; cDNA; 1056 BP.
XX
XX ADS46729;
AC
XX
XX 02-DEC-2004 (first entry)
DT
XX
XX Bacterial polynucleotide #1472.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
XX Bacteria.
OS
XX
XX US2003233675-A1.
PN
XX
XX 18-DEC-2003.
PD
XX
XX 20-FEB-2003; 2003US-00369493.
PF
XX
XX 21-FEB-2002; 2002US-0360039P.
PR
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI
XX
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT

```

PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
PS Claim 1; SEQ ID NO 25159; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1056 BP; 281 A; 188 C; 262 G; 325 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 13; Length 1056;
Best Local Similarity 94.4%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 413 ACGGCATCGTCAGTTGCG 396
|||||

RESULT 7
ID AB213613
XX AB213613 standard; DNA; 1137 BP.
XX
AC AB213613;
XX
DT 21-JAN-2003 (first entry)
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1418.
XX
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
XX Arabidopsis thaliana.
XX
XX WO200216655-A2.
XX
XX 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US026685.
XX
XX 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
XX (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
PI WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed and

PT producing plants with increased tolerance to these abiotic stresses.
XX
XX Claim 144; SEQ ID NO 1418; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1137 BP; 310 A; 212 C; 273 G; 342 T; 0 U; 0 Other;

Query Match 91.1%; Score 16.4; DB 6; Length 1137;
Best Local Similarity 94.4%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 532 ACGGCATCGTCAGTTGCG 549
|||||

RESULT 8
AAC33636
ID AAC33636 standard; DNA; 1301 BP.
XX
XX AAC33636;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 3772.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0130077P.
PR 19-APR-1999; 99US-0130449P.
PR 21-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142300P.
PR 09-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 26-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

 Query Match 91.1%; Score 16.4; DB 3; Length 1301;
 Best Local Similarity 94.4%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 ACGGCATCGTCAGTTGCG 18
 ||||| ||||| ||||| ||||| |||||
 Db 573 ACGGCATCGTCAGTTGCG 590

 RESULT 9
 ADT46111
 ID ADT46111 standard; cDNA; 1362 BP.
 XX
 AC ADT46111;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #20862.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 44549; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1362 BP; 242 A; 464 C; 423 G; 233 T; 0 U; 0 Other;

 Query Match 91.1%; Score 16.4; DB 13; Length 1362;
 Best Local Similarity 94.4%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 ACGGCATCGTCAGTTGCG 18
 ||||| ||||| ||||| ||||| |||||
 Db 416 ACGGCATCGTCAGTTGCG 433

 RESULT 10
 AAZ36119
 ID AAZ36119 standard; DNA; 16 BP.
 XX
 AC AAZ36119;
 XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Primer derived from a nucleic acid sequence specific to EHEC.
 XX
 KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
 KW I891; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
 KW PCR primer; probe; ss.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 PN WO9955908-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-FR001000.
 XX
 PR 28-APR-1998; 98FR-00005329.
 XX
 PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
 XX
 PI Frechon DTM, Laure FC, Thierry D;
 XX
 DR WPI; 2000-013443/01.
 XX
 PT New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.
 XX
 PS Claim 5; Page 27; 48pp; French.
 XX
 CC AAZ36103-27 represent fragments derived from nucleic acid sequences
 CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
 CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
 CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
 CC 95.8% homologous with I891 of E. coli (nucleotides 1-406 of AAZ36102).
 CC The second sequence (AAZ36102) is associated with the presence of
 CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes

CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
XX
SQ Sequence 16 BP; 2 A; 5 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 17
Dy 1 CGGCATCGTCAGTTGCG 16

RESULT 11
ACA40338
ID ACA40338 standard; DNA; 930 BP.
AC ACA40338;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #21995.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU36468.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 28208; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: the sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at:
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 930 BP; 136 A; 278 C; 355 G; 161 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 8; Length 930;
Best Local Similarity 94.1%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
Dy 828 CGGCATCGTCAGTTGCG 844

RESULT 12
ACA38540
ID ACA38540 standard; DNA; 936 BP.
XX
AC ACA38540;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #20197.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Mycobacterium bovis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU34670.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 26410; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 936 BP; 136 A; 281 C; 357 G; 162 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 8; Length 936;

Best Local Similarity 94.1%; Pred. No. 3.4e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

||||| |||||||

Db 837 CGGCATCGTCAGTTGCG 853

RESULT 13

ACA36906/c

ID ACA36906 standard; DNA; 1392 BP.

XX AC ACA36906;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #18563.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX *Listeria monocytogenes*.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU33036.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 24776; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1392 BP; 343 A; 214 C; 336 G; 499 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 8; Length 1392;

Best Local Similarity 94.1%; Pred. No. 3.5e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

||||| |||||||

Db 762 CGGCATCGTCAGTTGCG 746

RESULT 14

ACF70905/c

ID ACF70905 standard; DNA; 1563 BP.

XX AC ACF70905;

XX 20-NOV-2003 (first entry)

XX Photorhabdus luminescens nucleotide sequence #9372.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

XX detection; food; gene expression; plant; animal; microorganism; toxin;

XX antibiotic; biopesticide; virulence factor; disease model; plague;

XX whooping cough; gene; ds.

XX Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.


```

PR 07-FEB-2001; 2001FR-00001659.
XX (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A,
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
XX Claim 2; SEQ ID NO 9372; 1205pp; French.
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
XX SQ Sequence 1563 BP; 387 A; 363 C; 509 G; 304 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 1563;
Best Local Similarity 94.1%; Pred. No. 3.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 CGGCATCGTCAGTTGCG 18
Db 882 CGGCATCGTCAGTAGCG 866
RESULT 15
AAQ96251/c
ID AAQ96251 standard; DNA; 1632 BP.
XX
XX AC AAQ96251;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-DEC-1995 (first entry)
XX
XX Fructosyltransferase gene of A. diazotrophicus.
XX
XX Fructosyltransferase; FTase; fructooligosaccharide; sweetener; fructan;
KW db.
XX
XX Gluconacetobacter diazotrophicus.
XX
XX Key Location/Qualifiers
FH CDS 1..1632
FT /*tag= a
FT
XX
XX EF663442-A1.
XX
XX 19-JUL-1995.
XX

```

```

PF 22-DEC-1994; 94EP-00203737.
XX
PR 23-DEC-1993; 93CU-00000125.
XX
XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.
XX
XX Sosa JGA, Garcia LH, Gonzalez AC, Sosa GS;
XX
XX WPI; 1995-247529/33.
DR P-PSDB; AAR79142.
XX
XX New fructosyltransferase from Acetobacter diazotrophicus - for the
PT prodn. of fructo-oligosaccharide and fructan cpds. from sucrose, useful
XX e.g. as low calorie sweeteners.
XX
XX Claim 1; Page 10-11; 16pp; English.
XX
XX A genomic library of A. diazotrophicus SRT4 was produced in pPW12 and
CC transformed to a lev-neg. mutant of the same strain. Colonies that
CC recovered the mucous phenotype were selected; 2 recombinant cosmids each
CC contained the same 7.8 kb insert, and the FTase gene was localized to a
CC 2.3 kb fragment that was cloned into pUC18 to give pUC1823; the sequence
CC of the FTase gene in this plasmid is given in AAQ96251. Cloning allows
CC expression of recombinant FTase in E. coli and Pichia pastoris. (Updated
CC on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
XX SQ Sequence 1632 BP; 307 A; 551 C; 485 G; 289 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 2; Length 1632;
Best Local Similarity 94.1%; Pred. No. 3.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 CGGCATCGTCAGTTGCG 18
Db 231 CGGCATCGTCAGTTGCG 215
RESULT 16
AAD57635/c
ID AAD57635 standard; DNA; 1797 BP.
XX
XX AC AAD57635;
XX
XX 20-NOV-2003 (first entry)
DT
XX Rice disease resistance gene, P1PLC1.
DE
XX Rice; abiotic stress tolerance; pathogen resistance; disease resistance;
KW grain quality; nutritional content; plant yield; P1PLC1; plant; gene; ds.
XX
XX Oryza sativa.
XX
XX Key Location/Qualifiers
FH CDS 1..1797
FT /*tag= a
FT /product= "Rice P1PLC1 protein"
XX
XX WO2003048319-A2.
XX
XX 12-JUN-2003.
XX
XX 27-NOV-2002; 2002WO-US038359.
XX
XX 30-NOV-2001; 2001US-0334501P.
XX
XX (SYGN-) SYNGENTA PARTICIPATIONS AG.
XX
XX Sainz MB, Salmeron J, Weislo L;
XX
XX WPI; 2003-505288/47.
DR P-PSDB; AAE38267.
XX

```

PT New nucleic acid from *Oryza sativa*, useful for altering abiotic stress
PT tolerance, pathogen or disease resistance or the grain quality,
PT nutritional content or yield in a plant.
XX
XX Claim 2; Page 139-140; 223pp; English.
XX
XX The invention relates to nucleic acid molecules from rice encoding
CC proteins for abiotic stress tolerance, enhanced pathogen or disease
CC resistance and altered nutritional quality. The sequences of the
CC invention are useful for altering abiotic stress tolerance, pathogen or
CC disease resistance or the grain quality, nutritional content or yield in
CC a plant. The present sequence is rice disease resistance gene, *PIPLC1*
XX
XX Sequence 1797 BP; 447 A; 541 C; 470 G; 339 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 10; Length 1797;
Best Local Similarity 94.1%; Pred. No. 3.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CGGCATCGTCAGTTGCG 18
Db 859 CGGCATCGTCAGTTGCG 843
RESULT 17
AAS59613
ID AAS59613 standard; DNA; 22934 BP.
XX
AC AAS59613;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein encoding DNA #108.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX
XX 02-JUN-2000; 2000US-0208841P.
XX
XX 07-JUN-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Claim 1; SEQ ID NO 108; 1069pp; English.
XX
XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by *P. acnes*. The disorders include
CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of *P. acnes* in a

CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for *P. acnes* proteins. These antibodies can be
CC used to downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU60884-AAU61133 and AAU67652-AAU67654. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 22934 BP; 4591 A; 7273 C; 6479 G; 4591 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 22934;
Best Local Similarity 94.3%; Pred. No. 4.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CGGCATCGTCAGTTGCG 18
Db 11527 CGGCATCGTCAGTTGCG 11543
RESULT 18
ACF64542
ID ACF64542 standard; DNA; 22934 BP.
XX
AC ACF64542;
XX
DT 17-OCT-2003 (first entry)
XX
DE Propionibacterium acnes DNA contig sequence #108.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine; ds.
KW
OS Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallie-Douglas J;
XX
XX WPI; 2003-381789/36.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a *P. acnes* protein.
XX
XX Claim 1; SEQ ID NO 108; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM5624-ABM64536) and to
CC immunogenic fragments of *P. acnes* polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a *P. acnes*
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; and a vaccine composition (comprising *P. acnes* polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of *P. acnes* in a
 CC patient; and a method for inhibiting the development of *P. acnes* in a
 CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against *P. acnes*, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a *P. acnes* DNA contig which is specifically claimed
 CC in the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

XX SQ Sequence 22934 BP; 4591 A; 7273 C; 6479 G; 4591 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 8; Length 22934;
 Best Local Similarity 94.1%; Pred. No. 4.8e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTGGC 18
 Db 11527 CGGCATCGTCAGTGGC 11543
 |||||

RESULT 19

ADR20887/C

ID ADR20887 standard; DNA; 25860 BP.

XX AC ADR20887;

XX DT 07-OCT-2004 (first entry)

XX DE Photorhabdus luminescens genomic DNA encoding two insecticide proteins.
 XX KW insecticide protein; transgenic plant; insect-resistant plant; wheat;
 XX KM maize; gene; ds.

XX OS Photorhabdus luminescens subsp. laumondii.

XX FH Key Location/Qualifiers
 XX CDS 20872..21306
 XX FT /*tag= a
 XX FT /product= "Photorhabdus luminescens insecticide protein
 XX FT #1"
 XX FT /*tag= b
 XX FT /product= "Photorhabdus luminescens insecticide protein
 XX FT #2"

XX PN WO2003087377-A1.

XX XX 23-OCT-2003.

XX PD 17-APR-2003; 2003WO-FR001239.

XX PF 17-APR-2002; 2002FR-00004798.

XX PR (INSP) INST PASTEUR.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX XX WPI; 2003-833739/77.

XX DR P-PSDB; ADR20889, ADR20891.

XX PT New nucleic acid from Photorhabdus luminescens, useful for producing
 XX PT insecticidal polypeptides active against Lepidoptera, and for producing
 XX PT insect resistant transgenic plants.

XX PS Claim 22; SEQ ID NO 1; 44pp; French.

XX XX The invention comprises the amino acid and coding sequences of proteins

XX XX

CC from Photorhabdus luminescens which are active against insects. The
 CC proteins of the invention are toxic to Lepidoptera, specifically the
 CC genera Plutella, Heliothis, Helicoverpa, Spodoptera and Ostrinia. The DNA
 CC and protein sequences of the invention are useful in the preparation of
 CC transgenic, insect-resistant plants, specifically wheat and maize. The
 CC present Photorhabdus luminescens genomic DNA contains the coding sequence
 CC for the two insecticide proteins of the invention.

XX SQ Sequence 25860 BP; 7315 A; 5392 C; 6047 G; 7106 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 11; Length 25860;
 Best Local Similarity 94.1%; Pred. No. 4.8e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTGGC 18
 Db 24907 CGGCATCGTCAGTGGC 24891
 |||||

RESULT 20

AAI99682_05

Continuation (6 of 45) of AAI99682 from base 500001 (Mycobacterium tuberculosis strain H
 WP Sequence split into 45 fragments LOCUS AAI99682 Accession AAI99682

WP	Fragment Name	Begin	End
WP	AAI99682_00	1	110000
WP	AAI99682_01	100001	210000
WP	AAI99682_02	200001	310000
WP	AAI99682_03	300001	410000
WP	AAI99682_04	400001	510000
WP	AAI99682_05	500001	610000
WP	AAI99682_06	600001	710000
WP	AAI99682_07	700001	810000
WP	AAI99682_08	800001	910000
WP	AAI99682_09	900001	1010000
WP	AAI99682_10	1000001	1110000
WP	AAI99682_11	1100001	1210000
WP	AAI99682_12	1200001	1310000
WP	AAI99682_13	1300001	1410000
WP	AAI99682_14	1400001	1510000
WP	AAI99682_15	1500001	1610000
WP	AAI99682_16	1600001	1710000
WP	AAI99682_17	1700001	1810000
WP	AAI99682_18	1800001	1910000
WP	AAI99682_19	1900001	2010000
WP	AAI99682_20	2000001	2110000
WP	AAI99682_21	2100001	2210000
WP	AAI99682_22	2200001	2310000
WP	AAI99682_23	2300001	2410000
WP	AAI99682_24	2400001	2510000
WP	AAI99682_25	2500001	2610000
WP	AAI99682_26	2600001	2710000
WP	AAI99682_27	2700001	2810000
WP	AAI99682_28	2800001	2910000
WP	AAI99682_29	2900001	3010000
WP	AAI99682_30	3000001	3110000
WP	AAI99682_31	3100001	3210000
WP	AAI99682_32	3200001	3310000
WP	AAI99682_33	3300001	3410000
WP	AAI99682_34	3400001	3510000
WP	AAI99682_35	3500001	3610000
WP	AAI99682_36	3600001	3710000
WP	AAI99682_37	3700001	3810000
WP	AAI99682_38	3800001	3910000
WP	AAI99682_39	3900001	4010000
WP	AAI99682_40	4000001	4110000
WP	AAI99682_41	4100001	4210000
WP	AAI99682_42	4200001	4310000
WP	AAI99682_43	4300001	4410000
WP	AAI99682_44	4400001	4411529

Query Match 85.6%; Score 15.4; DB 4; Length 110000;
 Best Local Similarity 94.1%; Pred. No. 5.6e+02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
db 102682 CGGCATCGGCAGTTGCG 102698

RESULT 21
AAI99682_06
Continuation (7 of 45) of AAI99682 from base 600001 (Mycobacterium tuberculosis strain H37Rv) split into 45 fragments
Sequence accession AAI99682
Accession AAI99682

```

Query Match      85.6%; Score 15.4; DB 4; Length 110000;
Best Local Similarity 94.1%; Pred. No. 5.6e+02;
Matches 16: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

RESULT 22

AAI99683_05

Continuation (6 of 44) of AAI99683 from base 500001 (Mycobacterium tuberculosis strain H37Rv)

WP Sequence	split into 44 fragments	LOCUS AAI99683	Accession AAI99683
WP	Fragment Name	Begin	End
WP	AAI99683_00	1	110000
WP	AAI99683_01	100001	210000

WP	AA199683_02	200001	310000
WP	AA199683_03	300001	410000
WP	AA199683_04	400001	510000
WP	AA199683_05	500001	610000
WP	AA199683_06	600001	710000
WP	AA199683_07	700001	810000
WP	AA199683_08	800001	910000
WP	AA199683_09	900001	1010000
WP	AA199683_10	1000001	1110000
WP	AA199683_11	1100001	1210000
WP	AA199683_12	1200001	1310000
WP	AA199683_13	1300001	1410000
WP	AA199683_14	1400001	1510000
WP	AA199683_15	1500001	1610000
WP	AA199683_16	1600001	1710000
WP	AA199683_17	1700001	1810000
WP	AA199683_18	1800001	1910000
WP	AA199683_19	1900001	2010000
WP	AA199683_20	2000001	2110000
WP	AA199683_21	2100001	2210000
WP	AA199683_22	2200001	2310000
WP	AA199683_23	2300001	2410000
WP	AA199683_24	2400001	2510000
WP	AA199683_25	2500001	2610000
WP	AA199683_26	2600001	2710000
WP	AA199683_27	2700001	2810000
WP	AA199683_28	2800001	2910000
WP	AA199683_29	2900001	3010000
WP	AA199683_30	3000001	3110000
WP	AA199683_31	3100001	3210000
WP	AA199683_32	3200001	3310000
WP	AA199683_33	3300001	3410000
WP	AA199683_34	3400001	3510000
WP	AA199683_35	3500001	3610000
WP	AA199683_36	3600001	3710000
WP	AA199683_37	3700001	3810000
WP	AA199683_38	3800001	3910000
WP	AA199683_39	3900001	4010000
WP	AA199683_40	4000001	4110000
WP	AA199683_41	4100001	4210000
WP	AA199683_42	4200001	4310000
WP	AA199683_43	4300001	4403765

Query Match	85.6%	Score 15.4;	DB 4;	Length 110000;
Best Local Similarity	94.1%	Pred. No. 5.6e+02;		
Matches 16: Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

RESULT 23
AAI99683_06
Continuation (7 of 44) of AAI99683 from base 600001 (Mycobacterium tuberculosis strain H37Rv)
wp Sequence split into 44 fragments LOCUS AAI99683 Accession AAI99683

Query Match 85.6%; Score 15.4; DB 4; Length 110000;
Best Local Similarity 94.1%; Pred. No. 5.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

AAI99683_16 1600001 1710000
AAI99683_17 1700001 1810000
AAI99683_18 1800001 1910000
AAI99683_19 1900001 2010000
AAI99683_20 2000001 2110000
AAI99683_21 2100001 2210000
AAI99683_22 2200001 2310000
AAI99683_23 2300001 2410000
AAI99683_24 2400001 2510000
AAI99683_25 2500001 2610000
AAI99683_26 2600001 2710000
AAI99683_27 2700001 2810000
AAI99683_28 2800001 2910000
AAI99683_29 2900001 3010000
AAI99683_30 3000001 3110000
AAI99683_31 3100001 3210000
AAI99683_32 3200001 3310000
AAI99683_33 3300001 3410000
AAI99683_34 3400001 3510000
AAI99683_35 3500001 3610000
AAI99683_36 3600001 3710000
AAI99683_37 3700001 3810000
AAI99683_38 3800001 3910000
AAI99683_39 3900001 4010000
AAI99683_40 4000001 4110000
AAI99683_41 4100001 4210000
AAI99683_42 4200001 4310000
AAI99683_43 4300001 4403765

Query Match 85.6%; Score 15.4; DB 4; Length 110000;
Best Local Similarity 94.1%; Pred. No. 5.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CGGCATCGTCAGTTGCG 18
| | | | | | | | | | | | | | | | | |
Db 4042 CGGCATCGTCAGTTGCG 4058

RESULT 24

ABA03041_06/c
Continuation (7 of 30) of ABA03041 from base 600001 (Listeria monocytogenes EGD-e genome
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041

WP Fragment Name Begin End
ABA03041_00 1 110000
ABA03041_01 100001 210000
ABA03041_02 200001 310000
ABA03041_03 300001 410000
ABA03041_04 400001 510000
ABA03041_05 500001 610000
ABA03041_06 600001 710000
ABA03041_07 700001 810000
ABA03041_08 800001 910000
ABA03041_09 900001 1010000
ABA03041_10 1000001 1110000
ABA03041_11 1100001 1210000
ABA03041_12 1200001 1310000
ABA03041_13 1300001 1410000
ABA03041_14 1400001 1510000
ABA03041_15 1500001 1610000
ABA03041_16 1600001 1710000
ABA03041_17 1700001 1810000
ABA03041_18 1800001 1910000
ABA03041_19 1900001 2010000
ABA03041_20 2000001 2110000
ABA03041_21 2100001 2210000
ABA03041_22 2200001 2310000
ABA03041_23 2300001 2410000
ABA03041_24 2400001 2510000
ABA03041_25 2500001 2610000
ABA03041_26 2600001 2710000
ABA03041_27 2700001 2810000
ABA03041_28 2800001 2910000
ABA03041_29 2900001 2944528

Query Match 85.6%; Score 15.4; DB 6; Length 110000;
Best Local Similarity 94.1%; Pred. No. 5.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CGGCATCGTCAGTTGCG 18
| | | | | | | | | | | | | | | | | |
Db 87899 CGGCATCGTCAGTTGCG 87883

RESULT 25

ACF67367_40/c
Continuation (41 of 57) of ACF67367 from base 4000001 (Photorhabdus luminescens nucleot.
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP Fragment Name Begin End
ACF67367_00 1 110000
ACF67367_01 100001 210000
ACF67367_02 200001 310000
ACF67367_03 300001 410000
ACF67367_04 400001 510000
ACF67367_05 500001 610000
ACF67367_06 600001 710000
ACF67367_07 700001 810000
ACF67367_08 800001 910000
ACF67367_09 900001 1010000
ACF67367_10 1000001 1110000
ACF67367_11 1100001 1210000
ACF67367_12 1200001 1310000
ACF67367_13 1300001 1410000
ACF67367_14 1400001 1510000
ACF67367_15 1500001 1610000
ACF67367_16 1600001 1710000
ACF67367_17 1700001 1810000
ACF67367_18 1800001 1910000
ACF67367_19 1900001 2010000
ACF67367_20 2000001 2110000
ACF67367_21 2100001 2210000
ACF67367_22 2200001 2310000
ACF67367_23 2300001 2410000
ACF67367_24 2400001 2510000
ACF67367_25 2500001 2610000
ACF67367_26 2600001 2710000
ACF67367_27 2700001 2810000
ACF67367_28 2800001 2910000
ACF67367_29 2900001 3010000
ACF67367_30 3000001 3110000
ACF67367_31 3100001 3210000
ACF67367_32 3200001 3310000
ACF67367_33 3300001 3410000
ACF67367_34 3400001 3510000
ACF67367_35 3500001 3610000
ACF67367_36 3600001 3710000
ACF67367_37 3700001 3810000
ACF67367_38 3800001 3910000
ACF67367_39 3900001 4010000
ACF67367_40 4000001 4110000
ACF67367_41 4100001 4210000
ACF67367_42 4200001 4310000
ACF67367_43 4300001 4410000
ACF67367_44 4400001 4510000
ACF67367_45 4500001 4610000
ACF67367_46 4600001 4710000
ACF67367_47 4700001 4810000
ACF67367_48 4800001 4910000
ACF67367_49 4900001 5010000
ACF67367_50 5000001 5110000
ACF67367_51 5100001 5210000
ACF67367_52 5200001 5310000
ACF67367_53 5300001 5410000
ACF67367_54 5400001 5510000
ACF67367_55 5500001 5610000
ACF67367_56 5600001 5648894

Query Match 85.6%; Score 15.4; DB 10; Length 110000;
Best Local Similarity 94.1%; Pred. No. 5.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CGGCATCGTCAGTTGCG 18
|||||
Db 11517 CGGCATCGTCAGTAGCG 11501

Search completed: March 11, 2005, 04:20:04
Job time : 28.3327 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 7.29114 Seconds
(without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-20

Perfect score: 18

Sequence: 1 acgcgcatcgctgagtcg 18

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:**

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:**
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:**
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:**
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:**
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:**
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	15.4	85.6	1632	1	Sequence 1, Appli
C 2	15.4	85.6	1632	1	Sequence 1, Appli
C 3	15.4	85.6	4403785	3	Sequence 2, Appli
C 4	15.4	85.6	4411529	3	Sequence 1, Appli
C 5	15	83.3	206433	4	Sequence 12417, A
C 6	15	83.3	254778	4	Sequence 9380, Ap
C 7	14.8	82.2	38	4	Sequence 11861, A
C 8	14.8	82.2	38	4	Sequence 13564, A
C 9	14.8	82.2	38	4	Sequence 13564, A
C 10	14.8	82.2	219	4	Sequence 6160, Ap
C 11	14.8	82.2	510	4	Sequence 6082, Ap
C 12	14.8	82.2	627	4	Sequence 6472, Ap
C 13	14.8	82.2	633	4	Sequence 6472, Ap
C 14	14.8	82.2	645	4	Sequence 6472, Ap
C 15	14.8	82.2	654	4	Sequence 6002, Ap
C 16	14.8	82.2	915	4	Sequence 9133, Ap
C 17	14.8	82.2	993	4	Sequence 8950, Ap
C 18	14.8	82.2	1182	4	Sequence 6304, Ap
C 19	14.8	82.2	48994	4	Sequence 14091, A
C 20	14.4	80.0	354	4	Sequence 1808, Ap
C 21	14.4	80.0	1587	4	Sequence 6729, Ap
C 22	14.4	80.0	2328	4	Sequence 6729, Ap
C 23	14.4	80.0	2469	4	Sequence 6749, Ap
C 24	14.4	80.0	2511	4	Sequence 935, App
C 25	14.4	80.0	3099	4	Sequence 5065, Ap
C 26	14.4	80.0	4230	4	Sequence 5711, Ap
C 27	14.4	80.0	4941	4	Sequence 5741, Ap

28 14.4 80.0 28194 4 US-09-902-540-1250 Sequence 1250, Ap
 C 29 14 77.8 40085 4 US-08-311-731A-26 Sequence 26, Appl
 C 30 13.8 76.7 38 2 US-08-292-620A-2095 Sequence 2095, Ap
 C 31 13.8 76.7 38 2 US-08-292-620A-2311 Sequence 2311, Ap
 C 32 13.8 76.7 38 2 US-08-292-620A-2372 Sequence 2372, Ap
 C 33 13.8 76.7 38 3 US-09-071-845-2095 Sequence 2095, Ap
 C 34 13.8 76.7 38 3 US-09-071-845-2311 Sequence 2311, Ap
 C 35 13.8 76.7 38 3 US-09-071-845-2372 Sequence 2372, Ap
 C 36 13.8 76.7 38 4 US-09-371-772B-7372 Sequence 7372, Ap
 C 37 13.8 76.7 38 4 US-09-371-772B-7458 Sequence 7458, Ap
 C 38 13.8 76.7 38 4 US-09-371-772B-7707 Sequence 7707, Ap
 C 39 13.8 76.7 38 4 US-09-371-772B-7719 Sequence 7719, Ap
 C 40 13.8 76.7 38 4 US-09-371-772B-7838 Sequence 7838, Ap
 C 41 13.8 76.7 38 4 US-09-371-772B-8012 Sequence 8012, Ap
 C 42 13.8 76.7 38 4 US-09-371-772B-8521 Sequence 8521, Ap
 C 43 13.8 76.7 38 4 US-09-371-772B-8708 Sequence 8708, Ap
 C 44 13.8 76.7 38 4 US-09-371-772B-9156 Sequence 9156, Ap
 C 45 13.8 76.7 38 4 US-09-371-772B-9193 Sequence 9193, Ap

ALIGNMENTS

RESULT 1

US-08-362-232-1/c

; Sequence 1, Application US/08362232

; Patent No. 5641667

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Fructosyltransferase Enzyme, Method

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann & Baron

; STREET: 350 Jericho Turnpike

; CITY: Jericho

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11758

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 Mb

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Wordperfect 6.0 for DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/362,232

; FILING DATE: 22-December-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: CU 125/93

; FILING DATE: 23-December-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Baron, Ronald J.

; REGISTRATION NUMBER: 29,281

; REFERENCE/DOCKET NUMBER: 294-29

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550

; TELEFAX: (516) 822-3582

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1632 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

US-08-362-232-1

Query-Match 85.6%; Score 15.4; DB 1; Length 1632;
 Best Local Similarity 94.1%; Pred. No. 77;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTCG 18

||||| |||||||

```
Db      231 CGGCATGTCAGTTGCG 215

RESULT 2
US-08-814-196-1/c
; Sequence 1, Application US/09814196
; Patent No. 5731173
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Fructosyltransferase Enzyme, Method
; TITLE OF INVENTION: For its Production and DNA Encoding the Enzyme.
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.0 for DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,196
; FILING DATE: 10-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,232
; FILING DATE: 22-December-1994
; APPLICATION NUMBER: CU 125/93
; FILING DATE: 23-December-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1632 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-814-196-1

Query Match      85.6%; Score 15.4; DB 1; Length 1632;
Best Local Similarity 94.1%; Pred. No. 77;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CGGCATGTCAGTTGCG 18
      |||||
Db      231 CGGCATGTCAGTTGCG 215

RESULT 3
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13527

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match      85.6%; Score 15.4; DB 3; Length 4411529;
Best Local Similarity 94.1%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 CGGCATGTCAGTTGCG 18
      |||||
Db      602682 CGGCATCGGCAGTTGCG 602698

RESULT 5
US-09-949-016-13527/c
; Sequence 13527, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13527
```


; LENGTH: 206433
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13527

Query Match 83.3%; Score 15; DB 4; Length 206433;
Best Local Similarity 100.0%; Pred. No. 2.e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTT 15
|||||
Db 163151 ACGGCATCGTCAGTT 163137

RESULT 6
US-09-949-016-12417/c
; Sequence 12417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12417
; LENGTH: 254778
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12417

Query Match 83.3%; Score 15; DB 4; Length 254778;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTT 15
|||||
Db 163496 ACGGCATCGTCAGTT 163482

RESULT 7
US-09-371-772B-9380/c
; Sequence 9380, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MEHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371.772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9380
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-371-772B-9380

Query Match 82.2%; Score 14.8; DB 4; Length 38;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 21 ACGGCCTCATCAGTTGCG 4

RESULT 8
US-09-371-772B-11861/c
; Sequence 11861, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MEHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371.772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11861
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
; NAME/KEY: misc feature
; LOCATION: (31)-(31)
; OTHER INFORMATION: n stands for inosine
US-09-371-772B-11861

Query Match 82.2%; Score 14.8; DB 4; Length 38;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 21 ACGGCCTCATCAGTTGCG 4

RESULT 9
US-09-371-772B-13564/c
; Sequence 13564, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MEHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371.772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; ORGANISM: Artificial Sequence

```
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13564
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n stands for inosine
US-09-371-772B-13564

Query Match      82.2%; Score 14.8; DB 4; Length 38;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTGCG 18
Db      21  ACGGCCTCATCAGTTGCG 4

RESULT 10
US-09-248-796A-7462/c
; Sequence 7462, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7462
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7462

Query Match      82.2%; Score 14.8; DB 4; Length 219;
Best Local Similarity 88.9%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTGCG 18
Db      65  ACGGCATCGACAGTTGCG 48

RESULT 11
US-09-252-991A-6160/c
; Sequence 6160, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6160
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
```

```
US-09-252-991A-6160

Query Match      82.2%; Score 14.8; DB 4; Length 510;
Best Local Similarity 88.9%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTGCG 18
Db      378  ACGGCCTCGGCAGTTGCG 361

RESULT 12
US-09-252-991A-6082/c
; Sequence 6082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6082
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6082

Query Match      82.2%; Score 14.8; DB 4; Length 627;
Best Local Similarity 88.9%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTGCG 18
Db      565  ACGGCCTCGGCAGTTGCG 548

RESULT 13
US-09-252-991A-6472
; Sequence 6472, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6472
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6472

Query Match      82.2%; Score 14.8; DB 4; Length 633;
Best Local Similarity 88.9%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTGCG 18
Db      339  ACGGCCTCGGCAGTTGCG 356
```

```
RESULT 14
US-09-248-796A-7463
; Sequence 7463, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7463
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7463

Query Match      82.2%; Score 14.8; DB 4; Length 645;
Best Local Similarity 88.9%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ACGGCATCGTCAGTTGCG 18
      ||| ||||| ||||| |||
Db     154  ACGCCATCGACAGTTGCG 171

RESULT 15
US-09-252-991A-6002/c
; Sequence 6002, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6002
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6002

Query Match      82.2%; Score 14.8; DB 4; Length 654;
Best Local Similarity 88.9%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ACGGCATCGTCAGTTGCG 18
      ||| ||||| ||||| |||
Db     479  ACGGCCTCGGACAGTTGCG 462

RESULT 16
US-09-252-991A-9133
; Sequence 9133, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9133
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9133

Query Match      82.2%; Score 14.8; DB 4; Length 915;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ACGGCATCGTCAGTTGCG 18
      ||| ||||| ||||| |||
Db     827  ACGGCATCGTCAGCGCG 844

RESULT 17
US-09-252-991A-8950/c
; Sequence 8950, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8950
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8950

Query Match      82.2%; Score 14.8; DB 4; Length 993;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ACGGCATCGTCAGTTGCG 18
      ||| ||||| ||||| |||
Db     110  ACGGCATCGTCAGCGCG 93

RESULT 18
US-09-252-991A-6304
; Sequence 6304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6304
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6304
```

```
Query Match      82.2%; Score 14.8; DB 4; Length 1182;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
    ||||| ||||| ||||| |||||
Db 332 ACGGCTCGGAGTTGCG 349

RESULT 19
US-09-949-016-14091
; Sequence 14091, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14091
; LENGTH: 48994
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14091

Query Match      82.2%; Score 14.8; DB 4; Length 48994;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
    ||||| ||||| ||||| |||||
Db 30985 ACGGCATCGTCAGTTGCG 31002

RESULT 20
US-09-489-039A-1808
; Sequence 1808, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1808
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-1808

Query Match      80.0%; Score 14.4; DB 4; Length 354;
Best Local Similarity 93.8%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GGCATCGTCAGTTGCG 18
    ||||| ||||| ||||| |||||
Db 257 GGCATCGTCAATTGCG 272

RESULT 21
US-09-489-039A-6670/c
; Sequence 6670, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6670
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-6670

Query Match      80.0%; Score 14.4; DB 4; Length 1587;
Best Local Similarity 93.8%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GGCATCGTCAGTTGCG 18
    ||||| ||||| ||||| |||||
Db 810 GGCATCGTCAGTTGCG 795

RESULT 22
US-09-252-991A-5729
; Sequence 5729, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5729
; LENGTH: 2328
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-5729

Query Match      80.0%; Score 14.4; DB 4; Length 2328;
Best Local Similarity 93.8%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTTC 17
    ||||| ||||| ||||| |||||
Db 289 CGGCATCGTCAGTTTC 304

RESULT 23
US-09-489-039A-6749
; Sequence 6749, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
```

; SEQ ID NO 6749
; LENGTH: 2469
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6749

Query Match 80.0%; Score 14.4; DB 4; Length 2469;
Best Local Similarity 93.8%; Pred. No. 2.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GGCATCGTCAGTTGCG 18
|||
Db 1699 GGCATCGTCAGTTGCG 1714
|||

RESULT 24

US-09-023-655-935/c
; Sequence 935, Application US/09023655
; Patent No. 6607879

GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 935:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2511 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GENBANK

; CLONE: g1381141

US-09-023-655-935

Query Match 80.0%; Score 14.4; DB 4; Length 2511;
Best Local Similarity 93.8%; Pred. No. 2.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 17
|||
Db 1505 CGGCATCGTCAGTTGCG 1490
|||

RESULT 25

US-09-902-540-5065

; Sequence 5065, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 5065

; LENGTH: 3099

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-5065

Query Match

Best Local Similarity 80.0%; Score 14.4; DB 4; Length 3099;

Matches 15; Conservative 93.8%; Pred. No. 2.9e+02;

Mismatches 0; Indels 1; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 17

|||

Db 321 CGGCATCGTCAGTTGCG 336

Search completed: March 11, 2005, 13:13:31

Job time : 17.2911 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 89.924 Seconds
(without alignments)

1190.710 Million cell updates/sec

Title: US-09-674-277-20

Perfect score: 18

Sequence: 1 acggcatcgctgagttgag 18

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.4	91.1	664	17	US-10-338-110-22
C 2	16.4	91.1	1056	17	US-10-369-493-25159
C 3	16.4	91.1	1137	9	US-09-938-842A-1418
C 4	16.4	91.1	1137	11	US-09-938-842A-1418
C 5	16.4	91.1	1362	17	US-10-369-493-44549
C 6	15.4	85.6	25	19	US-10-719-900-375244
C 7	15.4	85.6	414	18	US-10-437-963-83252
C 8	15.4	85.6	767	18	US-10-425-115-101173
C 9	15.4	85.6	930	17	US-10-282-122A-28208
C 10	15.4	85.6	936	17	US-10-282-122A-26410
C 11	15.4	85.6	985	17	US-10-425-114-6662

C 12	15.4	85.6	1332	17	US-10-424-599-13630	Sequence 13630, A
C 13	15.4	85.6	1392	17	US-10-282-122A-24776	Sequence 24776, A
C 14	15.4	85.6	1797	18	US-10-491-733-23	Sequence 23, Appl
C 15	15.4	85.6	1848	18	US-10-437-963-7327	Sequence 7327, Ap
C 16	15.4	85.6	1945	18	US-10-425-115-101174	Sequence 101174, A
C 17	15.4	85.6	2090	17	US-10-425-114-13165	Sequence 13165, A
C 18	15.4	85.6	2209	18	US-10-437-963-7325	Sequence 7325, Ap
C 19	15	83.3	695	18	US-10-425-115-143802	Sequence 143802, A
C 20	14.8	82.2	35	9	US-09-864-785-3855	Sequence 3855, Ap
C 21	14.8	82.2	38	9	US-09-864-785-1207	Sequence 1207, Ap
C 22	14.8	82.2	38	10	US-09-730-289B-2380	Sequence 2380, Ap
C 23	14.8	82.2	38	10	US-09-730-289B-2694	Sequence 2694, Ap
C 24	14.8	82.2	38	10	US-09-877-478-3149	Sequence 3149, Ap
C 25	14.8	82.2	38	15	US-10-156-306-2050	Sequence 2050, Ap
C 26	14.8	82.2	38	16	US-10-230-006-1168	Sequence 1168, Ap
C 27	14.8	82.2	38	17	US-10-342-902-3149	Sequence 3149, Ap
C 28	14.8	82.2	38	17	US-10-138-674-11657	Sequence 11657, A
C 29	14.8	82.2	38	17	US-10-138-674-14138	Sequence 14138, A
C 30	14.8	82.2	38	17	US-10-138-674-15841	Sequence 15841, A
C 31	14.8	82.2	38	18	US-10-287-949A-11657	Sequence 11657, A
C 32	14.8	82.2	38	18	US-10-287-949A-14138	Sequence 14138, A
C 33	14.8	82.2	38	18	US-10-287-949A-15841	Sequence 15841, A
C 34	14.8	82.2	38	18	US-10-712-672-3048	Sequence 3048, Ap
C 35	14.8	82.2	38	18	US-10-669-841-8040	Sequence 8040, Ap
C 36	14.8	82.2	252	18	US-10-425-115-68648	Sequence 68648, A
C 37	14.8	82.2	284	18	US-10-425-115-109953	Sequence 109953, A
C 38	14.8	82.2	346	16	US-10-029-386-24609	Sequence 24609, A
C 39	14.8	82.2	354	18	US-10-425-115-14416	Sequence 14416, A
C 40	14.8	82.2	420	17	US-10-282-122A-29266	Sequence 29266, A
C 41	14.8	82.2	436	9	US-09-969-347-79	Sequence 79, Appl
C 42	14.8	82.2	477	18	US-10-767-701-21148	Sequence 21148, A
C 43	14.8	82.2	491	18	US-10-425-115-73894	Sequence 73894, A
C 44	14.8	82.2	515	18	US-10-767-701-17243	Sequence 17243, A
C 45	14.8	82.2	558	16	US-10-029-386-10906	Sequence 10906, A

ALIGNMENTS

RESULT 1

US-10-338-110-22/c
; Sequence 22, Application US/10338110
; Publication No. US20040023254A1
; GENERAL INFORMATION:
; APPLICANT: Fuhrmann, Jeffrey J.
; APPLICANT: Romesser, James A.
; TITLE OF INVENTION: A Method To Assess Quorum Sensing Potential Of Microbial
; TITLE OF INVENTION: Communities
; FILE REFERENCE: HER-0056
; CURRENT APPLICATION NUMBER: US/10338,110
; CURRENT FILING DATE: 2003-01-07
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Pseudomonas aureofaciens
US-10-338-110-22

Query Match 91.1%; Score 16.4; DB 17; Length 664;
Best Local Similarity 94.4%; Pred. No. 95;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

Db 419 ACGGCATCGTCAGTTGCG 402

RESULT 2

US-10-369-493-25159/c
; Sequence 25159, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25159
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-25159

Query Match          91.1%; Score 16.4; DB 17; Length 1056;
Best Local Similarity 94.4%; Pred. No. 97;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 413 ACGGCATCGCAGTTGCG 396

RESULT 3
US-09-938-842A-1418
; Sequence 1418, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1418
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1418

Query Match          91.1%; Score 16.4; DB 9; Length 1137;
Best Local Similarity 94.4%; Pred. No. 97;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 532 ACGGCATGTCAGTTGCG 549

RESULT 4
US-09-938-842A-1418
; Sequence 1418, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
```

```
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1418
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1418

Query Match          91.1%; Score 16.4; DB 11; Length 1137;
Best Local Similarity 94.4%; Pred. No. 97;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 532 ACGGCATGTCAGTTGCG 549

RESULT 5
US-10-369-493-44549
; Sequence 44549, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44549
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-44549

Query Match          91.1%; Score 16.4; DB 17; Length 1362;
Best Local Similarity 94.4%; Pred. No. 97;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 416 ACGGCATCGTCAGTTGCG 433

RESULT 6
US-10-719-900-375244
; Sequence 375244, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
```


; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 375244
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-375244

Query Match 85.6%; Score 15.4; DB 19; Length 25;
Best Local Similarity 94.1%; Pred. No. 2.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 4 CGGCATCGTCAGTTCCG 20
|||||

RESULT 7
US-10-437-963-83252/c
; Sequence 83252, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83252
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82601C.1
US-10-437-963-83252

Query Match 85.6%; Score 15.4; DB 18; Length 414;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGC 17
|||||
Db 394 ACAGCATCGTCAGTTGC 378
|||||

RESULT 8
US-10-425-115-101173/c
; Sequence 101173, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101173
; LENGTH: 767
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure

; LOCATION: (1)...(767)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_23783C.1
US-10-425-115-101173

Query Match 85.6%; Score 15.4; DB 18; Length 767;
Best Local Similarity 94.1%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 362 CGGCATCGTCAGTTGCG 346
|||||

RESULT 9
US-10-282-122A-28208
; Sequence 28208, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28208
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28208

Query Match 85.6%; Score 15.4; DB 17; Length 930;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 828 CGGCATCGTCAGTTGCG 844
|||||

```

RESULT 10
; Sequence 26410, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 78614
; SEQ ID NO 26410
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Mycobacterium bovis
US-10-282-122A-26410

Query Match      85.6%; Score 15.4; DB 17; Length 936;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2  CGGCATCGTCAGTTGCG 18
Db      837 CGGCATCGGCAGTTGCG 853

RESULT 11
US-10-425-114-6662
; Sequence 6662, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

```

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6662
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700574406_FLI
US-10-425-114-6662

Query Match      85.6%; Score 15.4; DB 17; Length 985;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2  CGGCATCGTCAGTTGCG 18
Db      373 CGGCATCGTCGTTGCG 389

RESULT 12
US-10-424-599-13630/c
; Sequence 13630, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 13630
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1332)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_112315C.1
US-10-424-599-13630

Query Match      85.6%; Score 15.4; DB 17; Length 1332;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  ACGGCATCGTCAGTTGCG 17
Db      852 ACGGCCTCGTCAGTTGC 836

RESULT 13
US-10-282-122A-24776/c
; Sequence 24776, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELTRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

PRIOR FILING DATE: 2003-02-20

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

PRIOR FILING DATE: 2003-03-21

RESULT 17

US-10-425-114-13165/c
; Sequence 13165, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13165
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-OSROLIB3475022E03_FLI
US-10-425-114-13165

Query Match 85.6%; Score 15.4; DB 17; Length 2090;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 958 CGGCATCGTCAGTGCG 942

RESULT 18

US-10-437-963-7325/c
; Sequence 7325, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 7325
; LENGTH: 2209
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13931C.1
US-10-437-963-7325

Query Match 85.6%; Score 15.4; DB 18; Length 2209;
Best Local Similarity 94.1%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 1006 CGGCATCGTCAGTGCG 990

RESULT 19

US-10-425-115-143802/c
; Sequence 143802, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 143802
; LENGTH: 695
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(695)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_62625C.1
US-10-425-115-143802

Query Match 83.3%; Score 15; DB 18; Length 695;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GCATCGTCAGTTGCG 18
|||||
Db 683 GCATCGTCAGTTGCG 669

RESULT 20

US-09-864-785-3855/c
; Sequence 3855, Application US/09864785
; Patent No. US20020177568A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Draper, Ken
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Levels of NF-Kappa B
; FILE REFERENCE: 400/022 (MBH00-812-D)
; CURRENT APPLICATION NUMBER: US/09/864,785
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 3929
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3855
; LENGTH: 35
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic Acid
; NAME/KEY: misc_feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: Phosphorothioate 3'-Internucleotide Linkage
; NAME/KEY: misc_feature
; LOCATION: (1)..(7)
; OTHER INFORMATION: 2'-O-Methyl
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: 2'-O-Methyl
; NAME/KEY: misc_feature
; LOCATION: (13)..(25)
; OTHER INFORMATION: 2'-O-Methyl
; NAME/KEY: misc_feature
; LOCATION: (27)..(28)
; OTHER INFORMATION: 2'-O-Methyl
; NAME/KEY: misc_feature

LOCATION: (30)..(34)
; OTHER INFORMATION: 2'-O-Methyl
; NAME/KEY: misc feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: 2'-deoxy-2'-C-Allyl
; NAME/KEY: misc feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: n stands for inosine
; NAME/KEY: misc feature
; LOCATION: (35)..(35)
; OTHER INFORMATION: n stands for inverted deoxyabasic derivative
; US-09-864-785-3855

Query Match 82.2%; Score 14.8; DB 9; Length 35;
Best Local Similarity 88.9%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 19 ACGGCCTCATCAGTTGCG 2

RESULT 21

US-09-864-785-1207/c
; Sequence 1207, Application US/09864785
; Patent No. US20020177568A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Draper, Ken
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to
; FILE REFERENCE: 400/022 (MBH00-812-D)
; CURRENT APPLICATION NUMBER: US/09/864,785
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 3929
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1207
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic Acid
; NAME/KEY: misc feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n stands for inosine
US-09-864-785-1207

Query Match 82.2%; Score 14.8; DB 9; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 21 ACGGCCTCATCAGTTGCG 4

RESULT 22

US-09-730-289B-2380/c
; Sequence 2380, Application US/09730289B
; Publication No. US20030050259A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for Treatment of Cardiac Disease
; FILE REFERENCE: MBH00-864-A (400/006)
; CURRENT APPLICATION NUMBER: US/09/730,289B
; CURRENT FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/169,100
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 3897

SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2380
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-730-289B-2380

Query Match 82.2%; Score 14.8; DB 10; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 21 ACGGCCTCATCAGTTGCG 4

RESULT 23

US-09-730-289B-2694/c
; Sequence 2694, Application US/09730289B
; Publication No. US20030050259A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for Treatment of Cardiac Disease
; FILE REFERENCE: MBH00-864-A (400/006)
; CURRENT APPLICATION NUMBER: US/09/730,289B
; CURRENT FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/169,100
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 3897
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2694
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
; NAME/KEY: misc feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n stands for inosine
US-09-730-289B-2694

Query Match 82.2%; Score 14.8; DB 10; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
Db 21 ACGGCCTCATCAGTTGCG 4

RESULT 24

US-09-877-478-3149/c
; Sequence 3149, Application US/09877478
; Publication No. US20030068301A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Draper, Kenneth
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; APPLICANT: Morrissey, Dave
; TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
; FILE REFERENCE: MBH00-845-H (400/029)
; CURRENT APPLICATION NUMBER: US/09/877,478
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: US 07/882,712
; PRIOR FILING DATE: 1992-05-14
; PRIOR APPLICATION NUMBER: US 09/531,025
; PRIOR FILING DATE: 2000-03-20

; PRIOR APPLICATION NUMBER: US 09/636,385
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: US 09/696,347
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 08/193,627
; PRIOR FILING DATE: 1994-02-07
; PRIOR APPLICATION NUMBER: US 08/433,993
; PRIOR FILING DATE: 1995-05-04
; PRIOR APPLICATION NUMBER: US 08/434,504
; PRIOR FILING DATE: 1995-05-04
; PRIOR APPLICATION NUMBER: US 09/436,430
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 6586
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3149
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
; US-09-877-478-3149

Query Match 82.2%; Score 14.8; DB 10; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 21 ACGGCCTCATCAGTTGCG 4

RESULT 25
US-10-156-306-2050/c
; Sequence 2050, Application US/10156306
; Publication No. US20030119017A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to
; TITLE OF INVENTION: Levels of IKK-Gamma and PKR
; FILE REFERENCE: MEHB01-664-A (400/050)
; CURRENT APPLICATION NUMBER: US/10/156,306
; CURRENT FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 8013
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2050
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31)-(31)
; OTHER INFORMATION: n stands for inosine
; US-10-156-306-2050

Query Match 82.2%; Score 14.8; DB 15; Length 38;
Best Local Similarity 88.9%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 21 ACGGCCTCATCAGTTGCG 4

Search completed: March 12, 2005, 00:25:27
Job time : 90.924 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:21:01 ; Search time 238.188 Seconds
(without alignment)
2876.537 Million cell updates/sec

Title: US-09-674-277-20
Perfect score: 18
Sequence: 1 acggcatcgctcagttgag 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hcc.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_g881.*
- 9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.4	91.1	326	CA548625	C0808A11-
C 2	16.4	91.1	457	B2516385	B2516385 BOMQES6TF
C 3	16.4	91.1	575	AV442266	AV442266 AV442266
C 4	16.4	91.1	643	BH540890	BH540890 BOHUT82TF
C 5	16.4	91.1	1184	CNS0A3E1	BX829195 Arabidops
C 6	16.4	91.1	1197	CNS0A2FG	BX826295 Arabidops
C 7	16.4	91.1	1414	CNS0A42C	BX826777 Arabidops
C 8	16	88.9	765	BH473819	BH473819 BOGHIO2TR
C 9	16	88.9	783	BH473894	BH473894 BOGHIO3TR
C 10	15.4	85.6	204	CG772409	CG772409 1123009G0
C 11	15.4	85.6	330	A1782868	A1782868 614005H02
C 12	15.4	85.6	401	AV403400	AV403400 AV403400
C 13	15.4	85.6	414	BH254127	BH254127 SALU_0160
C 14	15.4	85.6	422	CV163754	CV163754 rsmxl_00
C 15	15.4	85.6	438	CB486369	CB486369 omvkrtrCh0
C 16	15.4	85.6	456	BM270619	BM270619 sak1se02.
C 17	15.4	85.6	491	CB176223	CB176223 pl19e06.y
C 18	15.4	85.6	500	CB341518	CB341518 pk2e12.x
C 19	15.4	85.6	504	CB276633	CB276633 pl34e02.y
C 20	15.4	85.6	511	AG268681	AG268681 Cyanidios
C 21	15.4	85.6	525	CF642669	CF642669 D54_E07_F
C 22	15.4	85.6	526	CB176022	CB176022 pl17a03.y
C 23	15.4	85.6	528	BG653562	BG653562 sad58h09.
C 24	15.4	85.6	534	CB338954	CB338954 pl36h05.y

C 25	15.4	85.6	537	8	BH887961	BH887961 LB01768a.
C 26	15.4	85.6	544	6	CB174873	CB174873 pl14a06.y
C 27	15.4	85.6	545	7	CF775263	CF775263 NCESTqab1
C 28	15.4	85.6	549	8	BH887680	BH887680 LB01605a.
C 29	15.4	85.6	550	8	BH888276	BH888276 LB01957a.
C 30	15.4	85.6	571	6	CF024269	CF024269 QBS3C05.X
C 31	15.4	85.6	595	1	AJ716704	AJ716704 AJ716704
C 32	15.4	85.6	599	5	BP562650	BP562650 BP562650
C 33	15.4	85.6	605	8	BH521315	BH521315 BOGWM38TR
C 34	15.4	85.6	612	2	AW458601	AW458601 sh11c01.y
C 35	15.4	85.6	617	6	CA756273	CA756273 BR0300390
C 36	15.4	85.6	620	1	AV403156	AV403156 AV403156
C 37	15.4	85.6	620	6	CB190567	CB190567 pl28c08.y
C 38	15.4	85.6	623	4	BG761536	BG761536 602718763
C 39	15.4	85.6	625	8	CC117536	CC117536 NDL_71N1.
C 40	15.4	85.6	645	7	CF422438	CF422438 NCEST3d64
C 41	15.4	85.6	683	7	CO064902	CO064902 est_k_bre
C 42	15.4	85.6	688	6	CB001002	CB001002 S345H_H04
C 43	15.4	85.6	690	5	BP123165	BP123165 BP123165
C 44	15.4	85.6	718	1	AU293672	AU293672 AU293672
C 45	15.4	85.6	723	4	BM072893	BM072893 MEST55-D1

ALIGNMENTS

RESULT 1
CA548625/c 326 bp mRNA linear EST 19-NOV-2002
LOCUS C0808A11-5N NIA Mouse Blastocyst cDNA Library (Long) Mus musculus
DEFINITION C0808A11-5N NIA Mouse Blastocyst cDNA Library (Long) Mus musculus
ACCESSION CA548625
VERSION CA548625.1 GI:25092522
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 326)
AUTHORS Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Tanaka,T., Luo,A. and KO,M.S.H.

TITLE Systematic Analyses of NIA Mouse Blastocyst cDNA Library (Long)
JOURNAL Unpublished (2001)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdaa@leuon.grc.nia.nih.gov
Plate: C0808 row: A column: 11
Seq primer: M13 Reverse
High quality sequence stop: 326
POLYA=No.

FEATURES
source
1..326
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="C57BL/6J"
/db_xref="niaEST:C0808A11-5N"
/db_xref="taxon:10090"
/clone="NIA:C0808A11 IMAGE:30025738"
/tissue_type="Blastocyst"
/lab_host="DHI08"
/lab_stage="3.5-dpc"

/clone_lib="NIA Mouse Blastocyst cDNA Library (Long)"
NOTE: Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://leuon.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
extracted from a pool of 20 Blastocysts. Double-stranded
cDNAs were synthesized with an Oligo(dT) primer
[Invitrogen]:

5'-pGACTAGTCTAGATCGAGCGCGCCCTTTT-3' from 0.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPOR1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NTA)."

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 326;
Best Local Similarity 94.4%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 30 ACGGCATCGTCAGTTGCG 13

RESULT 2

BZ516385
LOCUS BZ516385 457 bp DNA linear GSS 16-DEC-2002
DEFINITION BOMQ56TF BO 2.3 KB Brassica oleracea genomic clone BOMQ56,
genomic survey sequence.

ACCESSION BZ516385
VERSION BZ516385.1 GI:27045849
KEYWORDS GSS.

ORGANISM

Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOMQ56TR
Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: T9

Class: sheared ends.

Location/Qualifiers

1. 457

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO100DH3"

/db_xref="taxon:3712"

/clone="BOMQ56"

/clone_lib="BO 2.3 KB"

/note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHO51 using BstXI linkers"

ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 457;
Best Local Similarity 94.4%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18
|||||
Db 394 ACGGCATCGTCAGTTGCG 411

RESULT 3

AV442266
LOCUS AV442266 575 bp mRNA linear EST 18-FEB-2004
DEFINITION AV442266 Arabidopsis thaliana above-ground organ two to six-week
Old Arabidopsis thaliana cDNA clone AP210H06_r 5', mRNA sequence.

ACCESSION

AV442266

VERSION

AV442266.1 GI:7612672

KEYWORDS

EST.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 575)

Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.

A large scale analysis of cDNA in Arabidopsis thaliana: Generation

of 12,028 non-redundant expressed sequence tags from normalized and

size-selected cDNA libraries

DNA Res. 7 (3), 175-180 (2000)

20363093

10907847

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yakusa 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

Location/Qualifiers

1. 575

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/clone="AP210H06_r"

/tissue_type="aboveground organs"

/dev_stage="two to six-week old"

/clone_lib="Arabidopsis thaliana above-ground organ two to

six-week old"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:

XhoI"

ORIGIN

Query Match 91.1%; Score 16.4; DB 1; Length 575;
Best Local Similarity 94.4%; Pred. No. 3.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

|||||

Db 535 ACGGCATCGTCAGTTGCG 552

RESULT 4

BH540890/c
LOCUS BH540890 643 bp DNA linear GSS 14-DEC-2001
DEFINITION BOHJU82TF BOHJ Brassica oleracea genomic clone BOHJU82, genomic
survey sequence.

ACCESSION

BH540890

VERSION

BH540890.1 GI:17788573

KEYWORDS

GSS.

SOURCE

Brassica oleracea

ORGANISM

Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

1 (bases 1 to 643)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Other GSSs: BOHJU82TR

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES

Location/Qualifiers

1. .643

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TOL000DH3"

/db_xref="taxon:3712"

/clone="BOHJ082"

/clone_lib="BOHJ"

/note="Vector: pHS01, Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS01 using BstXI linkers"

ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 643;
Best Local Similarity 94.4%; Pred. No. 3.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

Db 454 ACGGCATCGTCAGTTGCG 437

RESULT 5

CNS0A3E1

LOCUS

DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTSL802A12 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION

VERSION BX829195.1 GI:42461276

KEYWORDS HTC; GSLT cDNA.

SOURCE

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1184)

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

JOURNAL

REFERENCE 2 (bases 1 to 1184)

AUTHORS

Genoscope.

TITLE

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

Location/Qualifiers

1. .1184

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTSL802A12"

gene
/tissue_type="Silique"
/plasmid="pCMVSPORT.6"
complement(1..1184)
/gene="At4g36530"

ORIGIN

Query Match

Best Local Similarity 91.1%; Score 16.4; DB 3; Length 1184;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGCG 18

Db 537 ACGGCATCGTCAGTTGCG 554

RESULT 6

CNS0A2FG

LOCUS

DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFL12B03 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION

VERSION BX826295.1 GI:42459835

KEYWORDS HTC; GSLT cDNA.

SOURCE

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1197)

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

JOURNAL

REFERENCE 2 (bases 1 to 1197)

AUTHORS

Genoscope.

TITLE

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

Location/Qualifiers

1. .1197

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTFL12B03"

/tissue_type="Flowers and buds"

/plasmid="pCMVSPORT.6"

complement(1..1197)

/gene="At4g36530"

ORIGIN

Query Match

Best Local Similarity 91.1%; Score 16.4; DB 3; Length 1197;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1  ACGGCATCGTCAGTTGCG 18
        ||||| ||||| ||||| |||||
Db      553 ACGGCATTCGTCAGTTGCG 570

RESULT 7
CNS0A42C
LOCUS   1414 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            GSLTFB62D01 of Flowers and buds of strain col-0 of Arabidopsis
            thaliana (thale cress).
ACCESSION BX826777
VERSION   BX826777.1 GI:42462043
KEYWORDS  HTC; GSLT cDNA;
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana

REFERENCE
AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpetti,C., Schachter,V.,
            Temple,G., Caboche.M., Weissenbach,J. and Salanoubat,M.
TITLE     Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 1414)
GENOSCOPE
DIRECT SUBMISSION
SUBMITTED (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
        source
        1..1414
            Location/Qualifiers
                1..1414
                    /organism="Arabidopsis thaliana"
                    /mol_type="mRNA"
                    /strain="Col-0"
                    /db_xref="taxon:3702"
                    /clone="GSLTFB62D01"
                    /tissue_type="Flowers and buds"
                    /plasmid="pCMVSPORT 6"
                    complement(1..1414)
                    /gene="At4g36530"

gene
        91.1%; Score 16.4; DB 3; Length 1414;
        Best Local Similarity 94.4%; Pred. No. 4e+02;
        Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTGCG 18
        ||||| ||||| ||||| |||||
Db      698 ACGGCATTCGTCAGTTGCG 715

RESULT 8
BH438319
LOCUS   765 bp DNA linear GSS 12-DEC-2001
DEFINITION BOGHI02TR BOGH Brassica oleracea genomic clone BOGHI02, genomic
            survey sequence.

```

```

ACCESSION BH438319
VERSION   BH438319.1 GI:17624040
KEYWORDS  GSS.
SOURCE    Brassica oleracea
ORGANISM  Brassica oleracea

REFERENCE
AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE     Whole genome shotgun sequencing of Brassica oleracea
JOURNAL   Unpublished (2001)
COMMENT    TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
            Location/Qualifiers
                1..765
                    /organism="Brassica oleracea"
                    /mol_type="genomic DNA"
                    /strain="TO1000DH3"
                    /db_xref="taxon:3712"
                    /clone="BOGHI02"
                    /clone_lib="BOGH"
                    /notes="vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
                    genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN
Query Match      88.9%; Score 16; DB 8; Length 765;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ACGGCATCGTCAGTTG 16
        ||||| ||||| ||||| |||||
Db      313 ACGGCATCGTCAGTTG 328

RESULT 9
BH473894
LOCUS   783 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGUB33TR BOGU Brassica oleracea genomic clone BOGUB33, genomic
            survey sequence.
ACCESSION BH473894
VERSION   BH473894.1 GI:17682005
KEYWORDS  GSS.
SOURCE    Brassica oleracea
ORGANISM  Brassica oleracea

REFERENCE
AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE     Whole genome shotgun sequencing of Brassica oleracea
JOURNAL   Unpublished (2001)
COMMENT    Other GSSs: BOGUB33TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
            Location/Qualifiers
                1..783
                    /organism="Brassica oleracea"
                    /mol_type="genomic DNA"

FEATURES
        source
        1..783
            Location/Qualifiers
                1..783
                    /organism="Brassica oleracea"
                    /mol_type="genomic DNA"

```

QY	2	CGGCATCGTCAGTTGCG 18
Db	41	CGGCATCGCAGTTGCG 25
RESULT 11		
AI782868		
LOCUS	614005HO2.x2 614 - root cDNA library from Walbot Lab Zea mays cDNA,	
DEFINITION	mRNA sequence.	
ACCESSION	AI782868	330 bp mRNA linear EST 30-JUN-1999
VERSION	AI782868.1 GI:5296188	
KEYWORDS	EST.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
Eukaryota;	Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD		
ciade; Panicoidae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 330)	
AUTHORS	Walbot,V.	
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford	
JOURNAL	University	
COMMENT	Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 614005 row: H column: 02. Location/Qualifiers 1..330 /organism="Zea mays" /mol_type="mRNA" /cultivar="W23" /db_xref="taxon:4577" /tissue_type="root" /dev_stage="3-4 days old" /lab_host="XLOLR" ./clone_lib="614 - root cDNA library from Walbot Lab" /note="Organ: root; Vector: pBluescriptII SK+; Site_1: ECORI; Site_2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"	
FEATURES		
source		
ORIGIN		
Query Match	85.6%; Score 15.4; DB 1; Length 330;	
Best Local Similarity	94.1%; Pred.No. 1.3e+03;	
Matches 16;	Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	2	CGGCATCGTCAGTTGCG 18
Db	251	CGGCATCGTCGGTTGCG 267
RESULT 12		
AV403400		
LOCUS	AV403400 Bombyx mori middle silkland 5th-instar larva Bombyx mori	
DEFINITION	cDNA clone msgV0424 T3, mRNA sequence.	
ACCESSION	AV403400	401 bp mRNA linear EST 06-FEB-2000
VERSION	AV403400.1 GI:6907488	
KEYWORDS	EST.	
SOURCE	Bombyx mori (domestic silkworm)	
ORGANISM	Bombyx mori	
Eukaryota;	Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;		
Bombycoidea; Bombycidae; Bombyx.		
REFERENCE	1 (bases 1 to 401)	
AUTHORS	Mita,K., Morino,M., Shimada,T., Okano,K. and Maeda,S.	
TITLE	Bombyx mori cDNA	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: Mita K	

Genome Research Group
National Institute of Agrobiological Sciences
Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
Email: knita@nias.affrc.go.jp
method: uni-directional, sequence direction: sequenced from T3 primer
(5' -> 3')
Project= Silkorm Genome Program in MAFF, and Research for the
Future Program in JSPS. see 'SilkBase',
<http://www.ab.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.

FEATURES

source
1. .401
/organism="Bombyx mori"
/mol_type="mRNA"
/db_xref="taxon:7091"
/clone="msgV0424"
/sex="female/male mixed"
/tissue_type="middle silk gland"
/dev_stage="5th-instar larva"
/clone_lib="Bombyx mori middle silk gland 5th-instar larva"
/note="donated by Dr. Sehnal, Czech"

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 401;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18
|||||
Db 19 CGGCATCATCAGTTGCG 35

RESULT 13

BH254127 414 bp DNA linear GSS 28-NOV-2001
LOCUS SALU_016043 Arabidopsis thaliana TDNA insertion lines Arabidopsis
DEFINITION thaliana genomic clone SALU_016043, genomic survey sequence.
ACCESSION BH254127
VERSION BH254127.1 GI:17142638
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (chale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

Alonso,J.M., Leese,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmermann,J. and Ecker,J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

JOURNAL

Unpublished (2001)

COMMENT

Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 958 453 4100 x1752
Fax: 958 558 6379
Email: ecker@salk.edu

This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
At3g61340.

Class: TDNA tagged.

FEATURES

source

Location/Qualifiers
1. .414
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotYPE="Col-0"
/db_xref="taxon:3702"
/clone="SALK_016043"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 414;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGC 17
|||||
Db 5 ACGGCATCGTCAGTTGC 21

RESULT 14

CV163754/c

LOCUS

DEFINITION rsmxl_002413.y1.scf cDNA Library of Salvia miltiorrhiza Salvia
miltiorrhiza cDNA 5', mRNA sequence.

ACCESSION CV163754

VERSION CV163754.1 GI:51952894

KEYWORDS EST.

SOURCE Salvia miltiorrhiza

ORGANISM Salvia miltiorrhiza

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Lamiales; Lamiaceae; Nepetoideae; Nepeteae;
Salvia.

1 (bases 1 to 422)

Wang,Z., Yan,Y. and Tian,W.

Gene Discovery in Salvia miltiorrhiza Bge. as Revealed by EST

Sequencing

Unpublished (2004)

Contact: Zhezhi Wang, Yaping Yan, Wei Tian

Plant Cell Engineering

(Wang, Yan) College of Life Sciences, Shaanxi Normal University;

(Tian) Hangzhou Genomics Research and Development Institute

No. 199 of Chang'an South Road, Xi'an, Shaanxi, 710062, The

People's Republic of China

Tel: 86-29-85308352

Fax: 86-29-85303736

Email: zzwang@snnu.edu.cn

PCR Primers

FORWARD: M13 R

BACKWARD: M13 F

Seq primer: 3'-gtaccagtcagacaa-5' reverse primer M13.

Location/Qualifiers

1. .422

/organism="Salvia miltiorrhiza"

/mol_type="mRNA"

/db_xref="taxon:226208"

/sex="hermaphrodite"

/tissue_type="Whole Youth Plant"

/dev_stage="cultured about 1mon-3mon"

/clone_lib="cDNA Library of Salvia miltiorrhiza"

/note="Organ: Whole Youth Plant; Vector: pBluescript;

Site 1: EcORI; Site 2: XhoII; mRNA isolation used

PolyAtract(r) mRNA Isolation Systems (PROMEGA, CAT

No.25200) and pBluescript(r) XR cDNA Library Construction

Kit (Invitrogen, CAT No200455.) to construct the cDNA

library."

ORIGIN

Query Match 85.6%; Score 15.4; DB 7; Length 422;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGC 17
|||||
Db 172 ACAGCATCGTCAGTTGC 156

RESULT 15

CB486369
LOCUS
DEFINITION CB486369 438 bp mRNA linear EST 01-APR-2003
Oncorhynchus mykiss reproductive Oncorhynchus mykiss
CDNA, mRNA sequence.

ACCESSION
VERSION CB486369
KEYWORDS
SOURCE EST.

ORGANISM
Oncorhynchus mykiss (rainbow trout)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 438)

GRASP Consortium, Davidson, W.S., Koop, B.F. and
http://web.uvic.ca/cbr/grasp.

TITLE
A survey of Salmo salar transcripts from high complexity cDNA
libraries

JOURNAL
COMMENT Unpublished (2002)

Contact: Koop BF

Centre for Biomedical Research

University of Victoria

PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada

Tel: 250 472 4067

Fax: 250 472 4075

Email: bkoop@vic.ca

Centre for Biomedical Research, University of Victoria cDNA

preparation and sequencing: Roberto Alberto, Marianne

Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.

bioinformatics: Gordon D Brown.

Location/Qualifiers

1. 438

/organism="Oncorhynchus mykiss"

/mol_type="mRNA"

/strain="Karl Vierke"

/db_xref="taxon:8022"

/clone_lib="Oncorhynchus mykiss reproductive"

/note="Vector: pcr4topo; Library Creator: Kristian R von

Schalburg; Rainbow trout tissue contributors: Mountain

Trout Sales (Sooke, B.C.)"

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 438;

Best Local Similarity 94.1%; Pred. No. 1.3e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGC 17

|||||

289 AAGGCATCGTCAGTTGC 305

RESULT 16

BM270619/c

LOCUS

DEFINITION BM270619 456 bp mRNA linear EST 06-JUL-2004

sak15s02.y1 Gm-cl075 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-cl075-3267 5' similar to TR:Q9SJZ6 Q9SJZ6 AT2G232370 PROTEIN. ;

mRNA sequence.

ACCESSION

VERSION BM270619

KEYWORDS

SOURCE EST.

ORGANISM

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 456)

REFERENCE

AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,

Shanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. This clone is

available through: Biogenetic Services, 801 32nd Ave, Brookings, SD

57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Seq primer: -40RP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1. 456

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Jack"

/db_xref="taxon:3847"

/clones="SOYBEAN CLONE ID: Gm-cl075-3267"

/tissue_type="differentiating somatic embryos cultured on

MSM6AC"

/lab_host="DH10B"

/clone_lib="Gm-cl075"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from mRNA isolated

from differentiating somatic embryos cultured on MSM6AC.

The library was prepared using the Stratagene pBluescript

II SK(+) library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a

poly(dT) sequence with an XhoI restriction site. EcoRI

adaptors were ligated to the blunt-ended cDNA fragments

followed by XhoI digestion. The cDNA fragments were

directionally cloned into the EcoRI-XhoI restriction site

of the pBluescript vector. The ligated cDNA fragments

were transformed into E.coli Electromax DH10B host cells.

Tissue culture and library construction were performed by

Francoise Thibaud-Nissen and Anu Khana (Lila Vodkin lab,

University of Illinois)."

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 456;

Best Local Similarity 94.1%; Pred. No. 1.3e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGC 17

|||||

73 ACGGCCTCGTCAGTTGC 57

RESULT 17

CB176223/c

LOCUS

DEFINITION CB176223 491 bp mRNA linear EST 31-JAN-2003

pl19e06.y3 Ancylostoma ceylanicum adult Ancylostoma ceylanicum cDNA

5' similar to TR:O44145 O44145 C44B12.1 PROTEIN. [1]; mRNA

sequence.

ACCESSION

VERSION CB176223

KEYWORDS

SOURCE

ORGANISM

Ancylostoma ceylanicum

Ancylostoma ceylanicum

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;

Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.

1 (bases 1 to 491)

REFERENCE

AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,

Taagarishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,

Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

Harvey, N., Schuck, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library donated by John Hawdon of The George Washington University,
Washington DC (mtmjh@wumc.edu). Claire Murphy and Dr. James
McCarter of Washington University, GSC, St. Louis, MO mass excised
the pBluescript phagemid from the Lambda ZAP II library.

FEATURES

source

```
1. .491
/organism="Ancylostoma ceylanicum"
/mol_type="mRNA"
/db_xref="taxon:53326"
/dev_stage="adult"
/lab_host="SOLR (Stratagene)"
/clone_lib="Ancylostoma ceylanicum adult"
/note="Vector: pBluescript SK + excised from Lambda ZAP II
(Stratagene); Site 1: XhoI; Site 2: EcoRI; Lambda ZAP II
Library (99% recombinants, average insert size 1500bp,
amplified one time (10e6 pfu) donated by John Hawdon of
The George Washington University, Washington
DC (mtmjh@wumc.edu). Claire Murphy and Dr. James McCarter
of Washington University GSC, St. Louis, MO mass excised
the pBluescript phagemid from the Lambda ZAP II library."
```

ORIGIN

```
Query Match      85.6%; Score 15.4; DB 6; Length 491;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2  CGGCATCGTCAGTTGCG 18
      ||||| ||||| ||||| |||||
Db  185 CGGCATCATCAGTTGCG 169
```

RESULT 18

CA341518

LOCUS

```
DEFINITION      CA341518.1 500 bp mRNA linear EST 09-JAN-2003
3' similar to TR:O44145 O44145 C44B12.1 PROTEIN. [1] ; mRNA
sequence.
```

ACCESSION

CA341518

EST

KEYWORDS

SOURCE

```
Ancylostoma ceylanicum
Ancylostoma ceylanicum
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
```

REFERENCE

AUTHORS

```
McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
```

TITLE

JOURNAL

COMMENT

Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone will not be made available due to an unidentified
microbial contamination of the source material. Library donated by
John Hawdon of The George Washington University, Washington
DC (mtmjh@wumc.edu). Claire Murphy and Dr. James McCarter of

FEATURES

source

```
1. .500
/organism="Ancylostoma ceylanicum"
/mol_type="mRNA"
/db_xref="taxon:53326"
/dev_stage="adult"
/lab_host="SOLR (Stratagene)"
/clone_lib="Ancylostoma ceylanicum adult"
/note="Vector: pBluescript SK + excised from Lambda ZAP II
(Stratagene); Site 1: XhoI; Site 2: EcoRI; Lambda ZAP II
Library (99% recombinants, average insert size 1500bp,
amplified one time (10e6 pfu) donated by John Hawdon of
The George Washington University, Washington
DC (mtmjh@wumc.edu). Claire Murphy and Dr. James McCarter
of Washington University GSC, St. Louis, MO mass excised
the pBluescript phagemid from the Lambda ZAP II library."
```

ORIGIN

```
Query Match      85.6%; Score 15.4; DB 6; Length 500;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  2  CGGCATCGTCAGTTGCG 18
      ||||| ||||| ||||| |||||
Db  404 CGGCATCATCAGTTGCG 420
```

RESULT 19

CB276633/c

LOCUS

```
DEFINITION      CB276633.1 504 bp mRNA linear EST 25-FEB-2003
5' similar to TR:O44145 O44145 C44B12.1 PROTEIN. [1] ; mRNA
sequence.
```

ACCESSION

CB276633

EST

KEYWORDS

SOURCE

```
Ancylostoma ceylanicum
Ancylostoma ceylanicum
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
```

REFERENCE

AUTHORS

```
McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
```

TITLE

JOURNAL

COMMENT

Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library donated by John Hawdon of The George Washington University,
Washington DC (mtmjh@wumc.edu). Claire Murphy and Dr. James
McCarter of Washington University, GSC, St. Louis, MO mass excised
the pBluescript phagemid from the Lambda ZAP II library.
High quality sequence stop: 493.

FEATURES

source

```
1. .504
/organism="Ancylostoma ceylanicum"
/mol_type="mRNA"
/db_xref="taxon:53326"
/dev_stage="adult"
/lab_host="SOLR (Stratagene)"
```

/clone_lib="Ancylostoma ceylanicum adult"
 /note="Vector: pBluescript SK + excised from Lambda ZAP II
 (Stratagene); Site 1: XhoI; Site 2: EcoRI; Lambda ZAP II
 library (99% recombinants, average insert size 1500bp,
 amplified one time (10e6 pfu) donated by John Hawdon of
 The George Washington University, Washington
 DC(nmjhgwunc.edu). Claire Murphy and Dr. James McCarter
 of Washington University GSC, St. Louis, MO mass excised
 the pBluescript phagemid from the Lambda ZAP II library."

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 504;
 Best Local Similarity 94.1%; Pred. No. 1.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

Db 164 CGGCATCATCAGTTGCG 148

RESULT 20

AG268681/c

LOCUS

DEFINITION Cyanidioschyzon merolae genomic DNA, reverse end of BAC
 clone.H1D04, genomic survey sequence.

ACCESSION

AG268681

VERSION

AG268681.1

KEYWORDS

GSS.

SOURCE

ORGANISM

Eukaryota; Rhodophyta; Bangiophyceae; Cyanidiales; Cyanidiaceae;
 Cyanidioschyzon.

REFERENCE

AUTHORS

Maruyama, S., Misumi, O., Ishii, Y., Asakawa, S., Shimizu, A.,
 Sasaki, T., Matsuzaki, M., Shin-i, T., Nozaki, H., Kohara, Y.,
 Shimizu, N. and Kuroiwa, T.

TITLE

The Minimal Eukaryotic Ribosomal DNA Units in the Primitive Red

Alga Cyanidioschyzon merolae

JOURNAL

REFERENCE

AUTHORS

Submitted (18-SEP-2003) Tadasu Shin-i, National Institute of
 Genetics, Genome Biology Lab.; 1-111 Yata, Mishima, Shizuoka
 411-8540, Japan (E-mail: tshini@genes.nig.ac.jp, Tel: 81-55-981-6854,
 Fax: 81-55-981-6855)

FEATURES

source

Location/Qualifiers

1..511

/organism="Cyanidioschyzon merolae strain 10D"

/mol_type="genomic DNA"

/strain="10D"

/db_xref="taxon:280699"

/clone="H1D04"

/clone_lib="C. merolae BAC HindIII library"

ORIGIN

Query Match 85.6%; Score 15.4; DB 9; Length 511;

Best Local Similarity 94.1%; Pred. No. 1.3e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

Db 353 CAGCATCGTCAGTTGCG 337

RESULT 21

CF642669/c

LOCUS

DEFINITION D54_E07 Filamentous Forced Diploid Ustilago maydis cDNA 3', mRNA

sequence.

ACCESSION

CF642669

VERSION CF642669.1

KEYWORDS GI:37410264

SOURCE EST.

ORGANISM Ustilago maydis

Ustilago maydis

Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;

Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.

REFERENCE

AUTHORS

Nugent, K.G., Choffe, K. and Saville, B.J.

TITLE

Gene Expression during Ustilago maydis Diploid Filamentous Growth:

EST Library Creation and Analyses

JOURNAL

COMMENT

Fungal Genet. Biol. 41 (3), 349-360 (2004)

Contact: Barry J. Saville

Saville Lab

University of Toronto

3359 Mississauga Road North, Mississauga, ON, L5L 1C6, Canada

Tel: 905 569 4702

Fax: 905 828 3792

Email: bsaville@utoronto.ca

Plate: UTM-UM-D126/7-054-UTM row: 07 column: E

Seq primer: T7 Reverse (5' GAGTATACGACTCACTATAGG 3')

High quality sequence stop: 525.

Location/Qualifiers

1..525

/organism="Ustilago maydis"

/mol_type="mRNA"

/strain="FBD12"

/db_xref="taxon:5270"

/cell_type="Mycelia"

/dev_stage="Filamentous diploid"

/clone_lib="Filamentous Forced Diploid"

/note="Vector: pSport; mRNA was extracted from diploid

mycelia. A cDNA library was constructed and

unidirectionally cloned into pSPORT plasmid, with the use

of the Superscript II cDNA Library Construction Kit."

ORIGIN

Query Match 85.6%; Score 15.4; DB 7; Length 525;

Best Local Similarity 94.1%; Pred. No. 1.3e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

Db 114 CGGCATCGTCAGTTGCG 98

RESULT 22

CF642669/c

LOCUS

DEFINITION

Ustilago maydis cDNA 3', mRNA

sequence.

ACCESSION

CF642669

VERSION

CF642669.1

KEYWORDS

EST.

SOURCE

ORGANISM

Ancylostoma ceylanicum

Ancylostoma ceylanicum

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloida;

Ancylostomatidae; Ancylostomatinae; Ancylostominae; Ancylostoma.

REFERENCE

AUTHORS

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
 Wylie, R., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
 Teagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

CF642669.1

GI:37410264

EST.

Ustilago maydis

Ustilago maydis

Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;

Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.

REFERENCE

AUTHORS

Nugent, K.G., Choffe, K. and Saville, B.J.

TITLE

Gene Expression during Ustilago maydis Diploid Filamentous Growth:

EST Library Creation and Analyses

JOURNAL

COMMENT

Fungal Genet. Biol. 41 (3), 349-360 (2004)

Contact: Barry J. Saville

Saville Lab

University of Toronto

3359 Mississauga Road North, Mississauga, ON, L5L 1C6, Canada

Tel: 905 569 4702

Fax: 905 828 3792

Email: bsaville@utoronto.ca

Plate: UTM-UM-D126/7-054-UTM row: 07 column: E

Seq primer: T7 Reverse (5' GAGTATACGACTCACTATAGG 3')

High quality sequence stop: 525.

Location/Qualifiers

1..525

/organism="Ustilago maydis"

/mol_type="mRNA"

/strain="FBD12"

/db_xref="taxon:5270"

/cell_type="Mycelia"

/dev_stage="Filamentous diploid"

/clone_lib="Filamentous Forced Diploid"

/note="Vector: pSport; mRNA was extracted from diploid

mycelia. A cDNA library was constructed and

unidirectionally cloned into pSPORT plasmid, with the use

of the Superscript II cDNA Library Construction Kit."

ORIGIN

Query Match 85.6%; Score 15.4; DB 7; Length 525;

Best Local Similarity 94.1%; Pred. No. 1.3e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

Db 114 CGGCATCGTCAGTTGCG 98

CF642669

526 bp mRNA linear EST 31-JAN-2003

Ustilago maydis cDNA 3', mRNA

sequence.

ACCESSION

CF642669

VERSION

CF642669.1

KEYWORDS

EST.

SOURCE

ORGANISM

Ancylostoma ceylanicum

Ancylostoma ceylanicum

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloida;

Ancylostomatidae; Ancylostomatinae; Ancylostominae; Ancylostoma.

REFERENCE

AUTHORS

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
 Wylie, R., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
 Teagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library donated by John Hawdon of The George Washington University,
 Washington DC (mcjm@wumc.edu). Claire Murphy and Dr. James
 McCarter of Washington University, GSC, St. Louis, MO mass excised
 the pBluescript phagemid from the Lambda ZAP II library.
 High quality sequence stop: 505.

FEATURES

source

Location/Qualifiers
 1..526

/organism="Ancylostoma ceylanicum"
 /mol_type="mRNA"
 /db_xref="taxon:53326"
 /dev_stage="adult"
 /lab_host="SOLR (Stratagene)"
 /clone_lib="Ancylostoma ceylanicum adult"
 /note="Vector: pBluescript SK+ excised from Lambda ZAP II
 (Stratagene); Site 1: XhoI; Site 2: EcoRI; Lambda ZAP II
 Library (99% recombinants, average insert size 1500bp,
 amplified one time (10e6 pfu) donated by John Hawdon of
 The George Washington University, Washington
 DC (mcjm@wumc.edu). Claire Murphy and Dr. James McCarter
 of Washington University GSC, St. Louis, MO mass excised
 the pBluescript phagemid from the Lambda ZAP II library."

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 526;
 Best Local Similarity 94.1%; Pred. No. 1.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGGCATCGTCAGTTGCG 18

Db 184 CGGCATCATCAGTTGCG 168

RESULT 23

BG653562/c

LOCUS

DEFINITION BG653562 528 bp mRNA linear EST 22-JUL-2004
 sad58h09.y1 Gm-cl075 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-cl075-2297 5' similar to TR:Q9SUZ6 Q9SUZ6 AT2622370 PROTEIN.
 ;, mRNA sequence.

ACCESSION BG653562

VERSION BG653562.1 GI:13790971

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 528)

Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project

Unpublished (1999)

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
 clone is listed in the 'Other ESTs on clone' field. This clone is
 available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
 High quality sequence stop: 410.

FEATURES

source

Location/Qualifiers

1..528

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Jack"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl075-2297"

/tissue_type="differentiating somatic embryos cultured on
 MSM6AC"

/lab_host="DH10B"

/clone_lib="Gm-cl075"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from differentiating somatic embryos cultured on MSM6AC.
 The library was prepared using the Stratagene pBluescript
 II SK(+) library construction kit. Complementary DNA was
 synthesized from mRNA using a primer consisting of a
 poly(dT) sequence with an XhoI restriction site. EcoRI
 adaptors were ligated to the blunt-ended cDNA fragments
 followed by XhoI digestion. The cDNA fragments were
 directionally cloned into the EcoRI-XhoI restriction site
 of the pBluescript vector. The ligated cDNA fragments
 were transformed into E.coli ElectroMax DH10B host cells.
 Tissue culture and library construction were performed by
 Françoise Thibaud-Nissen and Anu Khana (Lila Vodkin lab,
 University of Illinois)."

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 528;
 Best Local Similarity 94.1%; Pred. No. 1.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACGGCATCGTCAGTTGC 17

Db 467 ACGGCCTCGTCAGTTGC 451

RESULT 24

CB338954/c

LOCUS

DEFINITION CB338954 534 bp mRNA linear EST 13-MAR-2003
 pl36h05.y1 Ancylostoma ceylanicum adult Ancylostoma ceylanicum cDNA
 5', similar to TR:O44145 O44145 C44B12.1 PROTEIN. [1]; mRNA
 sequence.

ACCESSION CB338954

VERSION CB338954.1 GI:28941939

KEYWORDS EST.

SOURCE Ancylostoma ceylanicum

ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.

REFERENCE 1 (bases 1 to 534)

AUTHORS

McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
 Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
 Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
 Tsagarisvilli,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
 Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
 Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

CONTACT: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library donated by John Hawdon of The George Washington University,
 Washington DC (mcjm@wumc.edu). Claire Murphy and Dr. James
 McCarter of Washington University, GSC, St. Louis, MO mass excised
 the pBluescript phagemid from the Lambda ZAP II library.
 High quality sequence stop: 494.

FEATURES

source

Location/Qualifiers

1..534

/organism="Ancylostoma ceylanicum"

/mol_type="mRNA"


```

/db_xref="taxon:53326"
/dev_stage="adult"
/lab_hosts="SOLR (Stratagene)"
/clone_lib="Ancylostoma ceylanicum adult"
/note="Vector: phluescript SK + excised from Lambda ZAP II
(Stratagene); Site 1: XhoI; Site 2: EcoRI; Lambda ZAP II
Library (99% recombinants, average insert size 1500bp,
amplified one time (10e6 pfu) donated by John Hawdon of
The George Washington University, Washington
DC(mtjmhgwmc.edu). Claire Murphy and Dr. James McCarter
of Washington University GSC, St. Louis, MO mass excised
the pBluescript phagemid from the Lambda ZAP II library."

```

ORIGIN

```

Query Match      85.6%; Score 15.4; DB 6; Length 534;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2  CCGCATCGTCAGTTGCG 18
      |||||
DB     194 CCGCATCATCAGTTGCG 178

```

RESULT 25

BH887961/c

LOCUS

```

DEFINITION      BH887961          537 bp      DNA          linear      GSS      07-AUG-2002
major genomic clone LB01768a major Friedlin BAC Library Leishmania

```

ACCESSION

BH887961

VERSION

GSS.

KEYWORDS

GSS.

SOURCE

Leishmania major

Leishmania major

Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;

Leishmania.

1 (bases 1 to 537)

Myler, P.J., Vogt, C., Munden, H., Robertson, L., Sisk, E.,

Fazelinia, G., Aggarwal, G., Nelson, S., Seyler, A., Worthey, E.,

Stuart, K., and Ragland, M.

Leishmania major Friedlin BAC End Sequences

Unpublished (2002)

Other GSSs: LB01768a.d.SP6.1

Contact: Myler PJ

Seattle Biomedical Research Institute

4 Nickerson Street, Seattle, WA 98109-1651, USA

Tel: 206 284-8846

Fax: 206 284-0313

Email: mylerpj@sbri.org

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..537

/organism="Leishmania major"

/mol_type="genomic DNA"

/strain="Friedlin"

/db_xref="taxon:5664"

/clone="LB01768a"

/lab_host="E. coli GeneHogs + Trfa"

/clone_lib="Leishmania major Friedlin BAC Library"

/note="Vector: pCG270; Site 1: HindIII; Genomic DNA from

Leishmania major Friedlin in agarose blocks was partially

digested with HindIII, size selected, and ligated with

HindIII-digested pCG270 vector DNA. 10368 clones were

picked and arrayed in 384- and 96-well plates. Library

construction and arraying was carried out by ResGen

Corporation and clones and filters are available from

them"

ORIGIN

```

Query Match      85.6%; Score 15.4; DB 8; Length 537;
Best Local Similarity 94.1%; Pred. No. 1.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1  ACCGCATCGTCAGTTGC 17
      |||||
DB     264 ACCGCATCGTCAGTCC 248

```

```

Search completed: March 11, 2005, 13:01:19
Job time : 240.188 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 06:55:29 ; Search time 7438.72 Seconds
(without alignments)
1190.710 Million cell updates/sec

Title: US-09-674-277-1
Perfect score: 1489
Sequence: 1 ctgcagctcgagatgaag.....ctggaaggagcctggtgac 1489

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936.4	62.9	2126	18	US-10-425-821-59
2	521	35.0	2214	17	US-10-282-122A-42350
3	394.6	26.5	2205	17	US-10-369-493-43360
4	382	25.7	1358	17	US-10-282-122A-24989
5	375	25.2	2139	17	US-10-369-493-34522
6	357.6	24.0	2238	9	US-09-884-889-7
7	357.6	24.0	2238	11	US-09-922-185-7
8	344.8	23.2	2145	17	US-10-369-493-33329
9	334.6	22.5	2208	17	US-10-369-493-40847
10	330.6	22.2	2163	17	US-10-369-493-32184
11	323.2	21.7	2271	17	US-10-369-493-35930

12	310.4	20.9	2205	17	US-10-369-493-44369
13	305.4	20.5	2181	17	US-10-369-493-47330
14	305.4	20.5	2181	17	US-10-282-122A-6882
15	303.4	20.4	2166	17	US-10-369-493-36422
16	302.2	20.3	2214	17	US-10-369-493-40706
17	298.4	20.0	978	17	US-10-282-122A-26464
18	298.4	20.0	2223	17	US-10-282-122A-28440
19	298.4	20.0	82993	15	US-10-080-170-645
20	298.4	20.0	82993	18	US-10-080-170-645
21	298.4	20.0	82993	18	US-10-080-170-645
22	297.6	20.0	2133	17	US-10-369-493-41662
23	297.2	20.0	2226	17	US-10-369-493-45100
24	293.6	19.7	2244	17	US-10-282-122A-13926
25	289.8	19.5	2181	17	US-10-282-122A-38893
26	288.2	19.4	2181	17	US-10-282-122A-39763
27	286.8	19.3	2265	17	US-10-369-493-26408
28	285.8	19.2	2175	17	US-10-282-122A-41114
29	284.8	19.1	2175	17	US-10-282-122A-23932
30	279.2	18.8	2094	17	US-10-282-122A-25855
31	276.8	18.6	2144	17	US-10-282-122A-20071
32	271.8	18.3	2154	17	US-10-282-122A-9084
33	259.4	17.4	2391	17	US-10-282-122A-11339
34	253.6	17.0	2202	17	US-10-369-493-40013
35	253.6	17.0	2255	17	US-10-369-493-39652
36	253.6	17.0	2265	17	US-10-369-493-37624
37	253.6	17.0	2267	17	US-10-369-493-39276
38	253	17.0	2196	17	US-10-369-493-33066
39	251.4	16.9	2346	17	US-10-369-493-41329
40	251.4	16.9	2731748	18	US-10-297-465A-1
41	250.4	16.8	2253	17	US-10-282-122A-32050
42	248.2	16.7	2198	17	US-10-369-493-33001
43	243.8	16.4	2268	17	US-10-282-122A-33671
44	241.4	16.2	772	17	US-10-369-493-33244
45	239.6	16.1	2244	17	US-10-369-493-26854

ALIGNMENTS

RESULT 1

US-10-425-821-59
; Sequence 59, Application US/10425821
; Publication No. US20040219530A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: HAREL, Jos,e
; APPLICANT: BEKAL, Sadjia
; TITLE OF INVENTION: ARRAY AND USES THEREOF
; FILE REFERENCE: 86369-3
; CURRENT APPLICATION NUMBER: US/10/425,821
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 2126
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-10-425-821-59

Query Match	62.9%	Score	936.4	DB	18	Length	2126
Best Local Similarity	95.9%	Pred. No.	4.3e+301	Mismatches	1	Indels	0
Matches	937	Conservative	0	Gaps	0		
Qy	552	CTTCCTGTTCTGATTTCTTGGCGCTATCGGGAGCTTTTCTACCGTGTAGCCGCTGAT	611				
Db	1	CTTCCTGTTCTGATTTCTTGGCGCTATCGGGAGCTTTTCTACCGTGTAGCCGCTGAT	60				
Qy	612	AAAAAGAGACTCAAAATTTCTACTATCCAGAAACACATGGATTTAACTCTCTGAGATTA	671				
Db	61	AAAAAGAGACTCAAAATTTCTACTATCCAGAAACACATGGATTTAACTCTCTGAGATTA	120				
Qy	672	CACAGCCCTGATCAAAATTCCTGGGGGCTGATTTTGTATATGCACAGATTTCAACAG	731				

Db 121 CACAGCCCTGAATCAAAATCCCTGGGGGCTGATTTTGGATTATGACCACAGATTTCAACAG 180
Qy 732 CTGATATGGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTGG 791
Db 181 CTGATATGGAGGCTCTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTGG 240
Qy 792 TGCCCTGGGATTTATGGTCATTATGGTCCTTTCTTTTATTCGTATGGCTTGGCAGGTGCC 851
Db 241 TGGCTGGGATTTATGGTCATTATGGTCCTTTCTTTTATTCGTATGGCTTGGCAGGTGCC 300
Qy 852 GGAACATACAGGACATATGATGGCGGGAGGCGCAGTGGTGGTCAGCAACGTTTGA 911
Db 301 GGAACATACAGGACATATGATGGCGGGAGGCGCAGTGGTGGTCAGCAACGTTTGA 360
Qy 912 CCGCTGAACAGCTGGCGGGATAACGTTTAACTGATTAAGCCGCTCGATTGCTGTGGCA 971
Db 361 CCGCTGAACAGCTGGCGGGATAACGTTTAACTGATTAAGCCGCTCGATTGCTGTGGCA 420
Qy 972 GTCAAGAAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTCTGACTGGTAAT 1031
Db 421 GTCAAGAAAAAATACGGCTCCAGTATTTCTTGGGGAGACCTGATGGTCTGACTGGTAAT 480
Qy 1032 GTTCCCTTTGAATCCATGGGATTTAAAAACGCTGGGATTTGCTGGCGGAAGAAGATGAC 1091
Db 481 GTTCCCTTTGAATCCATGGGATTTAAAAACGCTGGGATTTGCTGGCGGAAGAAGATGAC 540
Qy 1092 TGGAGTCTGGACCTGGGTATCTGGGGGCTGACAAAGCCCTTTCGACGATAACCGGGAT 1151
Db 541 TGGAGTCTGGACCTGGGTATCTGGGGGCTGACAAAGCCCTTTCGACGATAACCGGGAT 600
Qy 1152 AAAACGGGAACCTTCAGAAACCTCTTCCGCCACGACAGATGGGACTTTATTTATGTCAAT 1211
Db 601 AAAACGGGAACCTTCAGAAACCTCTTCCGCCACGACAGATGGGACTTTATTTATGTCAAT 660
Qy 1212 CCTGAAGCCCGGTGGAAAAACAGATCCTCTGGCTTCCGGGAAAGATATCAGGGAAGCT 1271
Db 661 CCTGAAGCCCGGTGGAAAAACAGATCCTCTGGCTTCCGGGAAAGATATCAGGGAAGCT 720
Qy 1272 TTTTCACGTATGGCCATGGATGATGAGAGACTGTGGCCCTGATCGGGAGGGCATACA 1331
Db 721 TTTTCACGTATGGCCATGGATGATGAGAGACTGTGGCCCTGATCGGGAGGGCATACA 780
Qy 1332 TTTGTAAGCACATGGTGACGCTCTCTGAAAAATGTTATTTGGCGAGGCGCTGATGGT 1391
Db 781 TTTGTAAGCACATGGTGACGCTCTCTGAAAAATGTTATTTGGCGAGGCGCTGATGGT 840
Qy 1392 GCACCTGTGGAGGACAGGACTGGGATGGAATAAATGTTGTTACAGGAAACGGCAAA 1451
Db 841 GCACCTGTGGAGGACAGGACTGGGATGGAATAAATGTTGTTACAGGAAACGGCAAA 900
Qy 1452 TATACCATCACAGTGGCTGGAGGAGCCTGGTGCAC 1489
Db 901 TATACCATCACAGTGGCTGGAGGAGCCTGGTGCAC 938

RESULT 2

US-10-282-122A-42350
; Sequence 42350, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42350
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: *Yersinia pestis*
US-10-282-122A-42350

Query Match 35.0%; Score 521; DB 17; Length 2214;
Best Local Similarity 71.7%; Pred. No. 3.2e-162;
Matches 683; Conservative 0; Mismatches 270; Indels 0; Gaps 0;
Qy 537 ATGATAAAAAAACTCTCTCTGTTCTGATTTCTTCTGGCGCTATCGGGAGCTTTTCTACC 596
Db 1 ATGTTAAAAAAATCTTTACCCGTACTAATAACTCTGCCATTTGATACATATACACCTACG 60
Qy 597 GCTGTAGCCGCTGATAAAAAAGAGACTCAAAATTTCTACTATATCCAGAAACACTGGATTTA 656
Db 61 GCTTGGCAGCAGAAGCACCCAGACTGATAGTTTCTTACTTACCCAAAAGCCTCGATCTC 120
Qy 657 ACTCTCTGAGTTTACACAGCCCTGATCAAACTCCCTGGGGGCTGATTTTGTATTATGTC 716
Db 121 TCTCCACTCCGTTTACATAATATAGAGTCCAATCTTACGAAAGGATTTTAATATGCT 180
Qy 717 ACCAGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAGATTTTCTGCACA 776
Db 181 CAACAGTTTAAACACTGGATCTTGAAGCAGTAAGAAGATATATAAACAAGTTCTTTACC 240
Qy 777 ACTTCCAGGATTTGGTCCCTCGCGGATTTATGGTCATTATGGTCTCTTTCTTTATTCGTATG 836
Db 241 ACATCAACAGACTGGTGGCCTGCTGATTATGGTAATATATGGTCCATTTCTTTATTCGTATG 300
Qy 837 GCTTGGCAGCTGCCGAAACATACAGCAGATATGATCGCCGGGAGCGCCAGTGGTGGT 896
Db 301 GCGTGGCATGTGGCGGAACCTACCGCATATATATGATGGCGGTGGTGGCGCGGG 360
Qy 897 CAGCAAGCTTTTGAACCGCTGAAACAGCTGGCCGAGATAACGTTAATCTGGATAAAGCCCGT 956
Db 361 CAGCAAGATTTGAGCCACTCAATAGCTGGCCAGATAACGCCAACCTTGTATAAAGCGGT 420
Qy 957 CGATTGCTGTGGCAGTCAAGAAAAATA CGGCTCCAGTATTTCTCTGGGGAGACTGATG 1016
Db 421 CGGCTTCTGTGGCCTATAAAAAAGAAATATGGCGCTAAAAATATCGTGGGGCGACTGATG 480
Qy 1017 GTCTGACTGGTAATGTTGCCCTTGAATCCATGGATTTAAAAACGCTGGGATTTGCTGGC 1076
Db 481 GTCTTACAGCAATGCGCGGCTCGAATCTATGGGGTTTAAAAACGCTGGGGTTTGCAGGA 540
Qy 1077 GGAAGAGAAGATGACTGGGAGTCCGAGCTGGTATATCTATCTGGGGGCTGACAAACAGCCTCTT 1136

Db 541 GCGCTGAGGATGATGGCAATCCGATCTGGTCTACTGGGGGAGGCAACAAGATGCTG 600
Qy 1137 GCAGATAACCGGATATAAAGCGGAACTTCAGAAACCTCTTGGCGCCACGAGATGGGA 1196
Db 601 TCTGATAACAGGGAATAAATGGCAAGCTACCCAAACCGCTGGCGGAGACAAATGGG 660
Qy 1197 CTTATTTATGTCATCTGTAAGGCGCGGTGGAACACAGATCTCTGGCTCCCGGAAA 1256
Db 661 TTGATTTACGTAAACCTGTAAGGACCAATGGTAAACACAGACCCGGTTCCTGCGCAAAA 720
Qy 1257 GATATCAGGGAAGCTTTTTCAGTATGGCATGGATGATGAGAGACTGTGGCCCTGATC 1316
Db 721 GATATTCGTGAAGCGTTTGGCAGGATGCGATGATGATGAGAGACCGTGGCATGAT 780
Qy 1317 GCGGAGGCGCATACATTTGGTAAAGCATGATGAGGCTCTCTGAAAAATGATTTGGC 1376
Db 781 GCGGAGGCGCATACCTTCGGTAAAGCGATGGCGAGCGTCTCTGAAAAATGCTGGGT 840
Qy 1377 GCAGGCGCTGATGGTGCACTCTGTGAGGAGGAGGAGCTGGATGGAAATTAATGCTGT 1436
Db 841 GCAGCAGCGGTGAAGCGGGGCTAGAACAAACAGGTTTAGGATGGCGGAATTAATGCTGT 900
Qy 1437 ACAGGAAACGGCAATATACCATCACCTGAGTGGCTGGAGGAGCTGTGTCGAC 1489
Db 901 TCAGGTATGTAGGACACGATACCAAGTGGTTGGAGGGGCGATGACAC 953

RESULT 3
US-10-369-493-43360
; Sequence 43360, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 43360
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Nitrosomonas europaea
US-10-369-493-43360

Query Match 26.5%; Score 394.6; DB 17; Length 2205;
Best Local Similarity 66.6%; Pred. No. 5,6e-120;
Matches 581; Conservative 0; Mismatches 289; Indels 3; Gaps 1;

Qy 620 GACTCAAAATTTCTACTATCCAGAAACACTGGATTAATCTCTCTGAGATATACACAGCCC 679
Db 81 GATGAACAGTTTGTGGTGGCGGATCACTGGATCTGAGACCGTTGGCGCAGAAATTCGGC 140
Qy 680 TGAATCAATCCCTGGGGGGCTGATTTGATATATGCCACAGATTTCAACAGCTGGATAT 739
Db 141 TGAGTCTAACCGTTGGGCAAGCGTTCAACTATGCCGAACAGTTTAAACCGCTCAACCT 200
Qy 740 GGAGCTCTGAAAAAGATATCAAGATTTTGTGACAACTTCCACAGGATTTGGTCCCTGC 799
Db 201 CAAGCGGTGAAGAGGATATCGGAAGTGTCTGATATCTTCCAGCGCTTGTGGCCCGC 260
Qy 800 GGATATGCTCATATGCTCTCTTTCTTATTTGATGCTTGGCAGCGTGGCGGAACATA 859
Db 261 GGACTACGGCAACTACGAGACCATTTGTTTCACTTCGATGCGCTGGCATAGTGGCGGTGA 320
Qy 860 CAGGACATATGATGGCGGGGAGGCGGCGAGTGGTGGTCAGCAACGTTTGTGAACCGCTGAA 919

Db 321 CCGCATTTTTCACGGAACGTGGTGGCGCTTCCGGCGGATAGCAACGCTTCGAACCGCTCAA 380
Qy 920 CAGCTGGCCGATTAACGTTAACTCTGGATAAAGCCCGTGCATTTGCTGTGGCCAGTCAAGAA 979
Db 381 CAGTTGGCCGATTAACGTTCAATCTGGACAAAGCGCGCGCTTGTATGGCCGATCAATA 440
Qy 980 AAAATACGGCTCCAGTATTTCTCTGGGAGAGACTGATGCTCTGATGTAATGTTGCCCT 1039
Db 441 GAAATACGCGCAGCAGCTTTCTCTGGGCTGATCTGATGCTGTGGCGCAATGTGGCGTT 500
Qy 1040 TGAATCCATGAGTATTTAAAGCGCTGGGATTTCTGGCGGAAGAGAGATGATGCGAGTC 1099
Db 501 GGAGGATATGGATTTCAAAACCATAGTTTTCGGGTGGGCGACCGGATGATTTGGGAAGC 560
Qy 1100 GGACCTGTATATCTGGGGCGCTGACAAACAGCCCTCTTTCAGATAAACCGG---GATAAAAA 1156
Db 561 CGAGATCGTCAACTGGGGAACAGAAAGGAATTTCTCGCCGATGAGCGCCCATGATAAAG 620
Qy 1157 CGGGAACCTTCAGAAACCTCTTGGCGGACCGACATGGGACTTATTTATGTCATCTGA 1216
Db 621 AGGTGAACCTGGCAGGCGCTCTGGCTGCGGTGCGATGCGGCTGATCTACGTCAACCCGA 680
Qy 1217 AGGCCCGCTGAAAAACAGATCTCTGGCTTCCGCGAAGATATCAGGGAAGCTTTTTC 1276
Db 681 AGGGCGGCGGTAAACCCCGATCCGTTGGCAGCTGCCAGGCATATCCCGAATCTCTCGG 740
Qy 1277 ACGTATGCGCATGATGAGGAGACTGTGGCCCTCTGATCCGGGAGGCGCATACATTTGG 1336
Db 741 CCGTATGGCGATGATGACGAAGAGACGGTGGCTCTGATCCCGGTGGCCATACCTTCGG 800
Qy 1337 TAAAGCATGCTGTCAGCGTCTCTGAAAAATGATTTGGCGCAGCGCTGATGTCGACC 1396
Db 801 CAAGCGCACGGGCGCATAGCCGGAAGATGGCTTGGCAAGGAACCGGCGCGGCTGG 860
Qy 1397 TGTGAGGAGCAGGACTGGGATGGAATAAATAATGTTGTCAGGAACCGCAATATAC 1456
Db 861 TATCAGAGAGCAAGGTTTGGTTGGGCAATAAATATGTCGCGGCTATGCGCTCGATAC 920
Qy 1457 CATCACCAGTGGCTGGAAGGAGCTGTGTCGAC 1489
Db 921 AGTCACCAGTGGTCTGGAAGGGGCGTGTGTCAC 953

RESULT 4
US-10-282-122A-24989
; Sequence 24989, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09015/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...2235
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-922-185-7

Query Match 24.0%; Score 357.6; DB 11; Length 2238;
Best Local Similarity 66.7%; Pred. No. 1.3e-107;
Matches 542; Conservative 0; Mismatches 264; Indels 6; Gaps 2;
Qy 683 ATCAATCCCTGGGGGCTGATTTTGTATGTCACACAGATTTCAACAGCTGGATATGGA 742
Db |||||
Qy 168 ATCGGACCAACACCCGGATTTTGACTATGCGCGAAGATTTAAGAAAGCTAGATCTGGC 227
Db |||||
Qy 743 GGCTCTGAATAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGGCGGA 802
Db |||||
Qy 228 AGCGTTAAAGAGACCTGGCAGCGCTAATGACAGATTCACAGACTGGTGGCCAGCAGA 287
Db |||||
Qy 803 TTATGGTCATPATGGTCTTTCTTTATTCGATATGGCTTGGCAGCGTCCGGAACATACAG 862
Db |||||
Qy 288 TTACGGTCATPATGGCCCTCTTTATACGATGCGCTGGCAGCAGCGCCGACCTACCG 347
Db |||||
Qy 863 GACATATGATGGCGGGAGCGGCAGTGGTGGTTCAGCAAGTTTGAACCGCTGACAG 922
Db |||||
Qy 348 TATCGGTGATGGCGTGGTGGCGTCTCGGCTCACAGCGCTTCGCGCTCTCAATAG 407
Db |||||
Qy 923 CTGCGCGGATAACGTTAATCTGGATAAGCCCGTCGATTGTGTGGCCAGTCAAGAAAAA 982
Db |||||
Qy 408 CTGCGCAGACATGCCAATCTGGATTAAGCAGCTTGCTTCTTTGGCCCATCAACAAA 467
Db |||||
Qy 983 ATACGGTCCAGTATTTCTCGGGAGACCTGATGGTCTGATCTGATATGTTGGCCCTTGA 1042
Db |||||
Qy 468 ATACGGTCGAAAAATCTCTGGCGGATCTAATGATACTCACAGAAACGTAGCTCTGGA 527
Db |||||
Qy 1043 ATCATGGGATTTAAACGCTGGGATTTGCTGGGAGAGAGATGACTGGGAGTC--- 1099
Db |||||
Qy 528 AACTATGGGCTTTAAACCTTTTGGTTTTCAGGTTGGCAGAGATGATGGGAGCCTGA 587
Db |||||
Qy 1100 GGACCTGTATCTAGGGGCTCTGACAAAGCCTCTTCAGATACCGGGATAAA---AA 1156
Db |||||
Qy 588 AGAAGATGATCTAGGGGAGCAGAAACCAATGGCTGGGAGACAGCGCTATGAAGTGA 647
Db |||||
Qy 1157 CGGGAACCTTCAGAAACCTTTGCGCGCACACAGATGGGACTATTTATGTCATCTCTGA 1216
Db |||||
Qy 648 CCGAGAGCTCGAAAAATCCCTGGGAGCGGTACAAATGGGACTCTATGATGTAACCCCGA 707
Db |||||
Qy 1217 AGGCCCCGGTGAACACAGATCTCTGCTTCGCGGAAGATATCAGGGAAGCTTTTC 1276
Db |||||
Qy 708 AGGACCCCAACGGCAGCAGACCCCTATCCCTGCTGCGCGTGATATTCGTGAGACTTTGG 767
Db |||||
Qy 1277 ACGTATGCCATGATGAGGAGCTGTGGCCCTGATCGCGGAGGGCATACATTTGG 1336
Db |||||
Qy 768 CCGAATGGCAATGATGACGAGAAACCGTGGCTCTCATAGGGGTGGACACACTTCGG 827
Db |||||
Qy 1337 TAAAGCAGATGGTCAGCGTCTCTCGAAAAATGATATGGCGCAGGGCCTGATGTTGAC 1396
Db |||||
Qy 828 AAAAACCCATGGTGTCCGATGCGGAGAAATATGTGGGCGCAGAGCGCTGCGCGCAGG 887
Db |||||

Qy 1397 TGTGGAGGACGAGGACTGGGATGGAAAAATAAATGTGTACAGGAAACGGCAAAATATAC 1456
Db |||||
Qy 888 TATTGAAGAAATGAGCTTGGGGTGGAAAAAACACACTACGGCACCGGACCGTGGGATAC 947
Db |||||
Qy 1457 CATCACAGTGGCTGGAGAGGACCTGGTGA 1488
Db |||||
Qy 948 CATCACAGTGGACTAGAAAGGCGCTGGACCA 979
Db |||||

RESULT 8

US-10-369-493-33329
; Sequence 33329, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33329
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Desulfitobacterium hafnense
US-10-369-493-33329

Query Match 23.2%; Score 344.8; DB 17; Length 2145;
Best Local Similarity 66.9%; Pred. No. 2.4e-103;
Matches 538; Conservative 0; Mismatches 257; Indels 9; Gaps 3;
Qy 687 AATCCCTGGGGGCTGATTTTGTATGTCACACAGATTTCAACAGCTGGATATGGAGCT 746
Db |||||
Qy 64 AATCCCATGATCGGACTTCAATATGCTGAAGATTTTCAGAACTTGACCTGGCAGCC 123
Db |||||
Qy 747 CTGAAAAAGATATCAAGATTTGCTGACAACTTCCAGGATTTGGTCCCTGGCGATAT 806
Db |||||
Qy 124 GTAAAGAAAGATCTTCTGCTTAAATGACCGGATTTCCAGGATTTGGTGGCTGGCGATTAC 183
Db |||||
Qy 807 GGTCAATTATGCTCTTCTTTATTCGATGCTGGCAGCTGGCAGCAATACAGGACA 866
Db |||||
Qy 184 GGCACACTACGGGCTCTCTTCATCCGGATGGCTTGGCAGCAGTGGGGAACATACCGTTG 243
Db |||||
Qy 867 TATGATCGCGGGAGCGCGCAGTGGTGGTCAAGCAAGTTTGAACCGCTGAAACAGCTGG 926
Db |||||
Qy 244 AACGACGACGGGGCGCGCGGAGAGCAACCCAGCGCTTTGCTCCCTCAACAGCTGG 303
Db |||||
Qy 927 CCGGATACCGTTAATCTGGATAAAGCCCGTCGATGCTGGCCAGTCAGAAAAAATAC 986
Db |||||
Qy 304 CCGGACAAATGTAAATCTGGATAGCGCCCGCTTGTCTCTGGCTTATTAACAGAAATAT 363
Db |||||
Qy 987 GGCTCAGTATTTCTTGGGAGACCTGATGGTCTGATGTTGCTTGGCTTGGCTTGAATCC 1046
Db |||||
Qy 364 GGAATAAATAATCTCTGGGCGGATCTGATGGTCTTGGCGGCAATTTGCTTGGATTC 423
Db |||||
Qy 1047 ATGGGATTTAAACGCTGGGATTTGCTGGGGAAGAGATGATCGGGATC---GGAC 1103
Db |||||
Qy 424 ATGGGCTTCAAGACTTTGGCTTCGCGCGCGCGCGGAGGATTTTGGAACTTCAGGAA 483
Db |||||
Qy 1104 CTGTATATCTGGGGCTCTGACAAACAGCCTTTCGAGATTAACCGGGATTAACAG---GG 1160
Db |||||
Qy 484 GATATTTATTGGGCTCTGAGAGAGAAATGGCTGGCGACACAGCGCTATTTCCGGGGATCGG 543
Db |||||
Qy 1161 AACTTCAGAAACCTCTTGGCGCCACCGAGATGGGACTTATTTATGTCATCTCGAAGGC 1220
Db |||||
Qy 544 GATCTTGAGAACCTCTCGCGCGAGTACAGATGGGCTGATTTATGTTAAACCCGGAAGGT 603
Db |||||

QY 759 ATCAAGATTGCTGACAACTTCCAGGATGTGGCCCTGGGATTAATGATTCATTATGGT 818
Db 178 CTGAAGCGCTGATGACCGATTCGAGTCCCTGGTGGCCGGGATTTCCGGCACTACGGC 237
QY 819 CCTTTCTTTATTCGTATGGCTTGGCAGCGTCCCGAACAATACAGGACATATGATGGCCGG 878
Db 238 CCGTTGTTTCGTCCGATGGCCCTGGCAGCGCAGGTACCTACCGCATCGGCGATGGCGGT 297
QY 879 GGAGCGCCAGTGTGTGTCAGCAAGCTTTTGAACCGCTGGAACAGCTGCCCGGATAACGTT 938
Db 298 GCGGTGTCGCGCTGGCCAGCAGCGTTTCGCGCCCAACAACAGCTGGCCGGAACAACGTC 357
QY 939 AATCTGGATAAAGCCCGTCGATGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGATTT 998
Db 358 AGCTGGACAAGGCACGACGCTCATCTGGCCGATCAAGCAAAATACGGCCGCAAGATC 417
QY 999 TCCTGGGAGACCTGATGGTCTGACTGTTAATGTTGCCCTTGAATCCATGGGATTTAAA 1058
Db 418 TCGTGGCCGACCTGATCGTTCTGACGGCAATGTTGCCCTGGAGTCGATGGGTTCAG 477
QY 1059 ACGCTGGGATTTGTCGGCGGAAGAGATGACTGGGAGTCGGACCTG---GTATACCTGG 1115
Db 478 ACCTTCGGATTCGGCGGAGGACGCGAGGATGCTATGAGCCGACGAGTCCGCTCTACTGG 537
QY 1116 GGGCTGACAAACAGCCTCTTGACAGATAACCGGGATA---AAAAAGGGAACCTTCAGAAA 1172
Db 538 GGCATGAAGCCGAGTGGCTGGCGGCAAGCGTTACGCGGTAAACCGGAACCTCGAGNAT 597
QY 1173 CCTCTTGGCGGACGACGATGGGACTTATTTATGTCAATCTGAAAGGCCCGGCTGGAAAA 1232
Db 598 CGCTGGCTCGGTGACAGTGGCCCTGATCTATGTAAATCCGGAAGGCCCAATGGCAAC 657
QY 1233 CCAAGTCTCTGGCTTCCGGAAGATATCAGGAGCTTTTTCAGTATGGCCATGGAT 1292
Db 658 CCGGACCCGGTTGCGCGCGCCATCGACATCCGCGAGACGTTCCCGCGCATGGCCATGAAC 717
QY 1293 GATGAGGAGCTGGCCCTGATCCGCGGAGGACATATTTGTTGTAAGACATGTTGCA 1352
Db 718 GACGAAGAAACCGTCCGCGCTGATCCGCGCGGTCTATGCTTCGGCAAGCATGGCGCC 777
QY 1353 GCGTCTCTGAAAAATGTTATGGCGCAGGGCTGATGTTGACCTCTGTGGAGGAGCAGGGA 1412
Db 778 G---GCCCGCATCGCACGTGGGCGCCGAGCTGAAGCCGCGGCTCGAGGAGCAGGGC 834
QY 1413 CTGGATGGAAAAATAAATGTTGTTACAGAAACGCAAAATATACCATCACCAGTGGCCCTG 1472
Db 835 CTTGGCTGGCGCAGCAGCTTTGGCACCGGCAAGGGCGGTGATGTCATCGGCAGTGGCCCTG 894
QY 1473 GAAGGAGCTGTGTCGAC 1489
Db 895 GAGGTCTCTGACCAC 911

RESULT 11
US-10-369-493-35930
; Sequence 35930, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35930
; LENGTH: 2271

; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35930
Query Match 21.7%; Score 323.2; DB 17; Length 2271;
Best Local Similarity 64.6%; Pred. No. 4.2e-96;
Matches 532; Conservative 0; Mismatches 283; Indels 9; Gaps 3;
QY 672 CACAGCCCTGAATCAATCCCTGGGGGCTGATTTTGATTTATGCACACAGATTTCAACAG 731
Db 190 CAGTCTCTCTGTCCGACCGGATGGGTGAGCGTTTCGACTATGCCGAGGAATTCAGAGC 249
QY 732 CTGATATGAGGCTCTGAAAAAAGATATCAAAAGATTTGCTGACAACTTCCAGGATGG 791
Db 250 CTGACCTCGACGCGCTCATCAGGACCTGCACGCGTGTATGACGGATTCGACGAGTGG 309
QY 792 TGGCTCGGATTAATGTCATTTATGTCCTTTCTTTATTTGTTATGTTGGCTTGCACGGTGGC 851
Db 310 TGGCGGCGCATTTCCGCCACTACGGCCCGCTGTTCATCCGGATGGCTTGGCACAGCGCA 369
QY 852 GBRACATACAGGACATATGATGCGGGGAGGCGCAGTGTGTGTCAGCAACGTTTGA 911
Db 370 GGCACCTTACCGCATTTGCCGACCGCGCGCGCGCTGCGAGCGGTTTCGCG 429
QY 912 CCGCTGAACAGCTGGCCGAGTAACTGATTAATCTGATTAAGCCCGTTCGATTTGCTGTGCCA 971
Db 430 CCGCTCAACAGCTGGCCGACATGTCACTCGACAAGCCCGCGCGCTTTTGTGGCG 489
QY 972 GTCAAGAAAAATACCGCTCCAGTATTTCTTGGGAGACCTGATGTTCTGACTGTTAAT 1031
Db 490 ATCAAGCAGAAATATGCGCGCAAGATCTCTTGGCGCGACCTTCTGATCTCACCGCAAC 549
QY 1032 GTTCCCTTGAATCCATGGGATTTAAACGCTGGGATTTGCTGCGGAAGAGAAGATGAC 1091
Db 550 GTCCGCTGGAATCGATGGGCTTCAGACCTTTTGGCTTTGCGCGCGCGCGCGCGCGCTC 609
QY 1092 TGGAGTTCGACCTG---GTATACTGGGGGCTTGACAACAGCCCTTTGTCAGATAACCGG 1148
Db 610 TGGGAGCTTGACGAGGAGCTTTACTGGGTCCCGAAGGCAAGTGGCTGGCGCAGAGCGC 669
QY 1149 GATA---AAACGGGAACTTTCAGAAACCTTTGCCCGCAGCAGATGGGACTTATTTAT 1205
Db 670 TACAGCGCGCACCTGACCTGCAAGACCCGCTCGCGCGCGGTGACAGATGGGCTCATCTAC 729
QY 1206 GTCAATCTCAGAGCCCGGTGGAAACCAAGATCCTCTGGCTTCCGCGAAAGATATCAGG 1265
Db 730 GTCAATTCGGAGGACCGAAGCGCAATCCGATCCGCTGGCCGCGCGCGCGCGATCAGG 789
QY 1266 GAAGCTTTTTCACGTATGGCCATGATGATGAGGAGACTGTGGCCCTGATCCGCGGAGG 1325
Db 790 GACACATTCGCGCTATGGCGATGAACGAGGAAACCGTAGCACCTATCGCGCGGC 849
QY 1326 CATACATTTGTTAAAGCACATGTTGTCAGCGTCTCTGAAATAATGTTTGGCGCAGGGCT 1385
Db 850 CATACGTTCCGCAAGACCCATGGT---GCGGGTGACGCGAGCCCTGGTGGGTGTGAGCGG 906
QY 1386 GATGTCACCTCTGGAGGAGCAGGACTGGATGGAATAATAATGTTGTTACAGGAAC 1445
Db 907 GAAGCGCTGATATCGAGCAGCAGGGCTTTGGCTGGGCGAGCAATTCGCGCACCGGCAAG 966
QY 1446 GCAAAATATACCATCACCAGTGGCTCGAAGAGGACCTGTTGTCGAC 1489
Db 967 GCGGTGACGCCATCGGACGCGGTCTGGAAGTCATTTGGACGAC 1010

RESULT 12
US-10-369-493-44269
; Sequence 44269, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

757	GGCATGAACGACGAAGAAACCGTGGCCTGATTGGCGGTGGTCATACGCTGGGTAAAAACC	816
1344	CATGGTGCACGCTCTCTGAAAAATGATTGGCGCAGCGGCTGATGCTGACCTGTGGAG	1403
817	CACGGTGCCG---GTCGACATCAAAATGTAGTCTGATCAGAAGCTGCACGGATTGAA	873
1404	GACGAGGACTGGATGGTAAAAATAATGTGTTACAGAAACGGCAAAATATACCATCACC	1463
874	GAACAAGTTTAGTTGGCGCAGCACTTACGGCAGCGCGTTGGCGCAGATGCCATTACC	933
1464	AGTGGCCTGAAGGAGCCTGGTC	1486
934	TCCTGGTCTGGAAGTAGTCTGGAC	956

RESULT 14

```

US-10-282-122A-6882
; Sequence 6882, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

	Query Match	20.5%	Score 305.4	DB 17	Length 2181
	Best Local Similarity	62.1%	Pred. No. 3.6e-90		
	Matches 499	Conservative 0	Mismatches 301	Indels 3	Gaps 1
Qy	684	TCAAAATCCCTGGGGGGCTGATTTTGTATTATCCACCAGATTTCAACAGCTGGATATGGAG	743		
Dh	157	TCTAACCCACTGGTGTAGGACTTTTGACTACCGCAAGAAGAAATTCAGCAAAATTAGATTACTATC	216		

RESULT 15
US-10-369-493-36422
; Sequence 36422, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36422

```

; LENGTH: 2166
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-10-369-493-36422

Query Match      20.4%; Score 303.4; DB 17; Length 2166;
Best Local Similarity 62.3%; Pred. No. 1.7e-89;
Matches 513; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

Qy 672 CACAGCCCTGAATCAAAATCCCTGGGGGGCTGATTTCGATTATGCACACAGATTTCAACAG 731
Db      |||||
Qy 100 CACAACTGGGTCTATATCCCTGGATAAGGGCTTTGACTATATCTGCGGGTTCATATAGC 159
Db      |||||
Qy 732 CTGATATGGAGGCTCTGAATAAAGATATCAAAGATTTGCTGCAAACTTCCCAGGATGG 791
Db      |||||
Qy 160 CTCGACTACTTCGGACTGAAGCGGGATCTGGAGGCACTCATGACAGACTCCAGGACTGG 219
Db      |||||
Qy 792 TGGCCCTGGGATATAGTTCATATATGGTCTCTTTCTTTATTTGCTATGGCTTGGCAGGTC 851
Db      |||||
Qy 220 TGGCCGGCCGACTTTGGTCACTATGGGGGACTCTTTTATCCGCATGGCTCGCACAGTGT 279
Db      |||||
Qy 852 GGAACATACAGGACATATGATGGCCGGGAGGCGCCAGTGTGTCAGCAACGTTTTGAA 911
Db      |||||
Qy 280 GGAACGTAATCGGTCTTTGACGTTCGGCGGGTGGCGACAGGGTCAGCAACGCTTCGCT 339
Db      |||||
Qy 912 CCGCTGAACAGCTCGCCGGATTAACGTTAAATCTGGATAAAGCCCTCGATTGCTGTGGCCA 971
Db      |||||
Qy 340 CCGCTCAACAGCTGGCCGGATACGTCAGCTTGGACAAGGCTGCTGCTCTCTTATGGCCC 399
Db      |||||
Qy 972 GTCAGAAAAAATACGGCTCAGTATTTCTTGGGGAGACCTGATGGTCTCTGACTGGAAT 1031
Db      |||||
Qy 400 ATCAAGCAGAAGTAGCGAAGCAAGATCTCATGGGCTGACTTGTGTGATTTCTCGCGGGAAT 459
Db      |||||
Qy 1032 GTTGCCCTTGAATCCATGGGATTTAAACCGTGGGATTTGCTGGCGAAAGAGAATGAC 1091
Db      |||||
Qy 460 GTGCCCTTGAATCAATGGGGTTTCAAGACTCTTTGGGTTTGCCTGGTGGCGCAAGTGATACC 519
Db      |||||
Qy 1092 TGGGAGTTCGGACCTGGGTATATCTGGGGGCTTGACAACAA-----GCCTCTTGCAGATAAC 1145
Db      |||||
Qy 520 TGGGAGACAGACAGTCTGGTCTTCTGGGAGGGCAGAGGAAGTGTGGGGTATGATGTC 579
Db      |||||
Qy 1146 CGGGATAAAAACGGGAAACTTTCAGAACTCTTTGCCGCCACGAGATGGGACTTATTTAT 1205
Db      |||||
Qy 580 CGCTACTTGAACGGAGAACTCGACAACCCGCTTGGCGCATCACACATGGGCTTATTTAC 639
Db      |||||
Qy 1206 GTCAATCTGAGGCCCGGTGGAAAAACAGATCCTCTGGGTTCCGCGAAAGATATCAGG 1265
Db      |||||
Qy 640 GTTAATCCAGAAGGACCAACAAGAACCCGACCCGGTTCTCGCGGCCAAGGATATCCGC 699
Db      |||||
Qy 1266 GAAGCTTTTTCAGTATGGCCATGGATGATGAGGAGACTGCGCCCTGATCGCGGAGGG 1325
Db      |||||
Qy 700 ATCACTTTGGTTCGAATGGCCATGAATGATGAGGAGACTGTGCGCCCTGATTTGCTGGTGA 759
Db      |||||
Qy 1326 CATACATTTGGTAAAGCACATGGTGCAGCGTCTCTGAAAAATGATTTATGGCGCAGGSCCT 1385
Db      |||||
Qy 760 CACAGTTTCGGAAGACGACGG--CGCGGGCCCTGCAACCCATCTCGGCAAGAACCA 816
Db      |||||
Qy 1386 GATGGTGCACTGTGGAGGACGAGGAGCTGGGATGGAAAAATAAATGTGTGACAGGAAC 1445
Db      |||||
Qy 817 CATGGTCGGGTATTGATGTTACAGGCCCTAGGCTGGGAGAGCGGCTTCGATCTGGGACC 876
Db      |||||
Qy 1446 GGCAAAATATACCATCACAGTGGCTCGAAGAGCCCTGGTCGA 1488
Db      |||||
Qy 877 GGGCGCATGCTATCACAGCGGTCTGGAGGTGATCTGGACCA 919
Db      |||||

```

RESULT 16
US-10-369-493-40706
; Sequence 40706, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

```

; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 40706
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Caulobacter crescentus
US-10-369-493-40706

```

Query Match	20.3%	Score 302.2;	DB 17;	Length 2214;
Best Local Similarity	62.0%;	Pred. No. 4.3e-89;		
Matches 513;	Conservative 0;	Mismatches 308;	Indels 6;	Gaps 2;
Qy	663	CTGAGATTACACAGCCCTGGAATCAAAATCCCTGGGGGCTGATTGTGATTATGCCACCAGA	722	
Db	136	CTGAACAGACAGCCCGCCGCTCAATCCGATGGCGAGCGTTCGACTACGCCGAGCG	195	
Qy	723	TTTCAACAGCTGGATATGGAGGCTCTGAAAAAGATATCAAGATTTCGTGCAAACTTCC	782	
Db	196	TTCAAGAGCCTCGACCTGGACGCGGTCGTACGCGACCTGCACGCCCTGATGCCGACAGC	255	
Qy	783	CAGGATTGGTCCCTGGCGATTATGTGTCATTATGTCTCTTTTATTCTGATGGCTTGG	842	
Db	256	CAGGAGTGGTGGCGGCTGACTTCGGGCCACTATGGCGGCTGTTCATTTCGCTGGCTTGG	315	
Qy	843	CACGGTCCCGGAACATACAGACATATGATGGCCGGGAGCGGCAGTGTGGTGCAGCAA	902	
Db	316	CACGGCCCGGACCTATCGCNTACCGACGGCGCGCGCGGGCGGGGCGGACGAG	375	
Qy	903	CGTTTTGAACCGGTGAACAGCTGGCCGGATPACGTTTAATCTGGATAAAGCCCTCGATTCG	962	
Db	376	CGCTTCGCCCCCTTGAAACAGCTGGCCGGACAAACCAATCTCGAACAGGCCCGCCGCTG	435	
Qy	963	CTGTGGCCAGTCAAGAAATAACGGCTCCAGTATTTCTCTGGGAGACCTGATGGTCTCTG	1022	
Db	436	CTGTGGCCGATCAAGACGAAGTAGTCGGCGCCAGCTGTCTGGGCGGACCTCTATGTGCTG	495	
Qy	1023	ACTGGTAAATGTGCCCCCTGGAATCCATGGGATTTAAACCGCTGGGATTTGCTGGCGGAAGA	1082	
Db	496	GTCCGCAACGTCGCCCTGGAGTCGATGGGCTTCAAGACCTTCGGCTTCGCGCGCGCGCG	555	
Qy	1083	GAGATGACTGGGAGTCGGACCTGGTATATCTGGGGGCTTGACAAAGCCCTCTTTCAGAT	1142	
Db	556	GCCGACAGTGGGAGCCCGAAGAGCTGTACTGGGGTCCGAAAGACAGTGGCTGGACGAC	615	
Qy	1143	AACCGGATAAAAACG--CGAAACTTCAGAAACCTCTTCGCCGACGAGATGGGACTT	1199	
Db	616	AAGCGCTACGCGCGAGCGTGTAGCTGGACTCGCCCTGGCGCGGCTCCAGATGGGCTG	675	
Qy	1200	ATTTATGTCAATCCTGAAGCCCGGTGGAAAAACAGATCCTCTGGCTTCGCGAAAGAT	1259	
Db	676	ATCTACGTCAATCCCGAAGCCCGAACCGCAACCCCGACCCGCTGGCTTCGGCGCGCGAC	735	
Qy	1260	ATCAGGGAAGCTTTTACGTTATGGCCATGATGATGAGAGACTGTGGCCCTGATCGCG	1319	
Db	736	ATTCGCGAGACCTTCGCCCCGATGGCCATGAATGACGAGGAGACCGTCGCCCTGATCGCC	795	
Qy	1320	GGAGGGCATATTTGGTAAAGCACATGGTGCAGCGTCTCTCGAAAAATGTATTGGCGCA	1379	
Db	796	GGCGGCCACACTTCGCAAGGCTCACGG---CGCGGGCAGCCCTCCCTCGTGGGCGTT	852	
Qy	1380	GGGCTGATGGTGCACTGTGGAGGAGCAGGAGACTGGGATGGAAAAATAAATGTGTGTA	1439	
Db	853	GAGCCGGAAGCGGCGCGATCTGAGGCCCAAGGCTTTGGTGGGCGAGCAAGCAGGCAGC	912	

```
QY 1440 GGAAACGGCAATATACCATCACCAGTGGCTCGAAGAGCCCTGGTC 1486
    |||||
Db 913 GGCAGGGCCCGACGCCATCACCGCGGCGCGGAGTGATCTGGAC 959
    |||||

RESULT 17
US-10-282-122A-26464
; Sequence 26464, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26464
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Mycobacterium bovis
US-10-282-122A-26464

Query Match 20.0%; Score 298.4; DB 17; Length 978;
Best Local Similarity 62.4%; Pred. No. 4.8e-88;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGGCTGATTTTGAATTATGCCACCATGATTTCAACAGCTGGATATGAGGCTC 747
    |||||
Db 122 ACCGATGGTGGCGGCTTCGACTATCCGCGGAGGTGCGGCACCATCGACGTTGACGCC 181
    |||||

QY 748 TGAAGAAAGATACAAAGATTGCTGACACTTCCAGGATTTGGTCCCTGCGGATTATG 807
    |||||
Db 182 TGACGCGGCATCGAGGAAGTGATGACCACTCGCAGCGGTGGTGGCCGCCGACTACG 241
    |||||

QY 808 GTCAATTATGTCCTTTCTTTATTCGTATGCTTGGCACGGTCCGGAACATACAGGACAT 867
    |||||
Db 242 GCACATACGGCCGCTGTTTATCCGATGGGTGGCACGCTGCCGCGACCTACCGCATCC 301
    |||||

QY 868 ATGATGCGCGGGAGGCGCCAGTGGTGGTCAGCAACGTTTGAACCGCTGAACAGCTGGC 927
    |||||
```

```
Db 302 ACGACGCGCGCGCGCGCGCGCGCGCGCATGCGAGCGGTTTCGCGCGCTTAAACAGCTGGC 361
QY 928 CGGATAACGTTAATCTGGATAAAGCCGCTCGATTGCTGTGGCCAGTCAAGAAAAATACG 987
    |||||
Db 362 CCGACAACCGCCAGCTTGGACAAGCGCGCGCGCTGCTGTGGCCGGTCAAGAGAGTACG 421
    |||||

QY 988 GCTCCAGTATTTCTCTGGGAGACCTGATGGTCTCTGACTGGTAAATGTTGCCCTTGAATCCA 1047
    |||||
Db 422 GCAAGAAGCTCTCATGGCGGACCTGATTGTTTTCGCGCGCAACTGCGCGCTGGAATCGA 481
    |||||

QY 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGGAAGAGATGACTGGGAGTCCGACCTGG 1107
    |||||
Db 482 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGGCCGGGTTCGACCATGCGGAGCCCGATGAGG 541
    |||||

QY 1108 TATACTGGGGGCTGTGACAAACAGCCCTCTTGCAGATAACCGGGATA--AAACCGGAAC 1164
    |||||
Db 542 TCTATTGGGGCAAGGAAGCCACTTGGCTCGCGATGAGCGTTACGGGTTAAGCGGGATC 601
    |||||

QY 1165 TTCAGAAACCTCTTTCGCCGCACGCGAGATGGGACTTATTTATGTCAATCTCTGAAGGCCCG 1224
    |||||
Db 602 TGGAGAACCCTGCGCGCGGTGCGAGATGGGCTGATCTACGTGAACCCCGAGGGCGCA 661
    |||||

QY 1225 GTGGAAACAGATCTCTTGGCTTCGCGGAAGATATCAGGGAAGCTTTTTCAGTATGG 1284
    |||||
Db 662 ACGCAACCCGACCCCATGGCGCGCGTTCGACATTCGCGAGACGTTTCGCGCGCATGG 721
    |||||

QY 1285 CCATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGCGCATATTTGGTAAAGCAC 1344
    |||||
Db 722 CCATGAACGACGTTCGAACACGCGCGCTGATCTCGCGGTTCACATTTTCGTTAAGACC 781
    |||||

QY 1345 ATGCTGACGCGTCTCCTGAAAAATGTTTGGCGCAGGGCCTGATGTTGTCACCTGTGGAGG 1404
    |||||
Db 782 ATGGCGCG--GCCCGCGGATCTGTGTCGCGCCGAAACCCGAGGCTGCTCGCTGGAGC 838
    |||||

QY 1405 AGCAGGAGCTGGGATCGAAAAATAAATGTGTACAGAAACGCAATATACCATCACA 1464
    |||||
Db 839 AGATGGGCTTGGGCTGGAAGAGCTCGTATGGCAACCGGAACCGGTAAGGACCGCATCACA 898
    |||||

QY 1465 GTGGCCTGGAAGGAGCTGTGCA 1488
    |||||
Db 899 GCGGCATCGAGGTCTGATGACCA 922
    |||||

RESULT 18
US-10-282-122A-28440
; Sequence 28440, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
```

```

; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28440
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28440

Query Match      20.0%; Score 298.4; DB 17; Length 2223;
Best Local Similarity 62.4%; Pred. No. 8.1e-88;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

Qy      688 ATCCCTGGGGCGTGATTTTGATATGACCAACAGATTTCACACGTGTGATGAGGCTC 747
Db      167 ACCCGATGGGTGCGCGGTTCGACTATGCGCGGAGGTGCGACCATCGACGTTGACGCC 226

Qy      748 TGA AAAAAGATATCAAAGATTGCTGCACACTCCACGAGTTGTCGCTGCCGATTATG 807
Db      227 TGACCGGGACATCGAGAAGTAGTACCACCTCGCAGCCGTGTGTGCGCCGCCACTACG 286

Qy      808 GTCAATTATGTCCTTTCTTTATTCGTATGGCTTGGCACCGTGC CGGAACATACAGGACAT 867
Db      287 GCCACTACGGGCCGCTGTTATCCGATGGGTGCGACGCTGCCGGCACCTACCGCATCC 346

Qy      868 ATGATGCGCGGGAGGCCACAGTGGTGTGACGAACCGTTTGAACCGCTGAACAGCTGGC 927
Db      347 ACGACGGCCGCGCGCGCGCGCGGCGCGCATGACAGCGGTTCCGCGCCGCTTAACAGCTGGC 406

Qy      928 CGGNATACGTTAATCTGGATAAAGCCCGTCGATTGCTGTGGCCAGTCAAGAAAAAATACG 987
Db      407 CCGACAACGCCAGCTTGGAACAAGCGCGCGCGGCTGCTGTGCGCCGCTCAAGAAGAAGTACG 466

Qy      988 GCTCCAGTATTTCTCGGGAGACCTGATGTGTCCTGACTGGTGAATGTTGCCCTTCAATCCA 1047
Db      467 GCAAGAGCTCTCATGSGCGGACCTGATGTTTTCGCGGCACTGCGCGCTGGAATCGA 526

Qy      1048 TGGGATTTAAACGCTCGGATTTGCTCGCGAAGAGAAGATGACTGGGAGTCGACCTGG 1107
Db      527 TGGGCTTCAAGACGTTCCGGTTCCGTTCCGCGGGTCGACCACTGGGAGGCCGATGAGG 586

Qy      1108 TATACTGGGGCCGTGACAAACAGCCTCTTGACAGATAACCGGGATA---AAACGGGAAC 1164
Db      587 TCTATTGGGCAAGGAAGCCACCTGGGCTCGCGGATGAGCGTTACAGCGGTGAAGCGGGATC 646

Qy      1165 TTCAGAAACCTTTGCGCCACAGATGGGACTTTATTATGTCAATCTCGAAGGCCCGC 1224
Db      647 TGGGAACCCGCTGCGCGCGGTGCAATGAGGCTGATCTACGTGAACCCGGAGGGGCCGA 706

Qy      1225 GTGAAACACAGATCCTCTGCTTCCGCGAAGATATCAGGGAAGCTTTTTTTCAGTATGG 1284
Db      707 ACGGCAACCCGGACCCCATGSCCGCGCGGTTCGACATTCGCGAGACGTTTCGGCGCATGG 766

Qy      1285 CCATGGATGATGAGGAGACTGTGGCCCTGATTCGCGGAGGGGCAATCATTTGGTAAAGCAC 1344
Db      767 CCATGAAACGAGCTGGAACACGCGCGCTGATTCGTTCGGCGGTTCACACTTTCGGTAAAGACC 826

Qy      1345 ATGTCGACGCTCTCCTGMAAAATGTAATGCGCGAGGGCCGTGATGTTGTCACCTCTGTGAGG 1404
Db      827 ATGCGCGCG---GCCCGCGCGATCTGGTTCGCGCCCGAACCCGAGGCTGCTCCGCTGGAGC 883

Qy      1405 AGCAGGACCTGGGATGAAAAATAAATGTGTGTACAGGAAACGGCAAAATATACCATCACCA 1464

```



```
QY 1285 CCATGGATGATGAGAGACTGTGGCCCTGTATCGCGGAGGCGATACATTTTGTAAAGCAC 1344
|||
Db 39732 CCATGAACGACGTGAAACACGCGGCTGTATCTGCGCGGTACACACTTTCGTGAAGACCC 39673
|||
QY 1345 ATGTGTGACGCTCTCTGAAATAATGTATTTGCGGAGGCGCTGTATGTTGTCACCTGTGAGG 1404
|||
Db 39672 ATGGCGCG---GCCCGCGCATCTGGTTCGCGCCGGAACCGGAGGCTGCTCGCGTGGAGC 39616
|||
QY 1405 AGCAGGACTGGATGGAATAATAATGTGTACAGGAACCGGAATATACCATCACCA 1464
|||
Db 39615 AGATGGGCTTGGGCTTGAAGAGCTCGTATGCGACCGGAACCGGTAAGACGCGATCACCA 39556
|||
QY 1465 GTGGCCTTGAAGGAGCCTGTGTCGA 1488
|||
Db 39555 GCGGCATCGAGGTCTGTATGGACGA 39532
|||

RESULT 20
US-10-080-170-645/c
; Sequence 645, Application US/10080170
; Publication No. US2004012122A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495-0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-08-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 645
; LENGTH: 82993
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-645

Query Match 20.0%; Score 298.4; DB 18; Length 82993;
Best Local Similarity 62.4%; Pred. No. 8.3e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGCTGATTTTGTATTCGATCCAGGATTTTCAACAGCTGGATATGAGGCTC 747
|||
Db 40332 ACCCGATGGGTGGCGCGTTCGACTATGCCGGAGGTGCGACCATCGACGTGACGCCC 40273
|||
QY 748 TGAATAAAGATATCAAGATTTGCTGCAAACTTCCAGGATTTGTCGCCCTCGGATTATG 807
|||
Db 40272 TGACGCGGACATCGAGGAGTGATGACCACTCGCAGCCGTGTGGCCCGGACTACG 40213
|||
QY 808 GTCAATTATGTCCTTTTATTCGATGCTGCGCAGGATTTGTCGCCCTCGGATTATG 867
|||
Db 40212 GCCACTACGGCGCTGTTTATCCGGATGGCGTGGCAGCTCGCGGACCTACCGCATCC 40153
|||
QY 868 ATGATGCGCGGAGGCGCCAGTGTGTGACAACTTTTGAACCGCTGAACAGCTGCC 927
|||
Db 40152 ACGACGCGCGCGCGCGCGCGGCGCATGACGCGGTTCGCGCGCTTAAACAGTGGC 40093
|||
QY 928 CGGATAACGTTAACTCGATATAAGCCCTCTTGCAGATAAACCGGATA---AAAAACGGAAAC 987
|||
Db 40092 CCGACNACCCAGCTTTGACNAGGCGCGCGGCTGTGTGCGGTCGAGNAGATACG 40033
|||
QY 988 GCTCCAGTATTTCTTGGGAGACCTGATGGTCCTGACTGGTAATGTTGCCCTTGAATCCA 1047
|||
Db 40032 GCAAGAAGCTCTCATGGCGGACCTGATTTGTTTTCGCGGCAACTGCGCGCTGGAATCGA 39973
|||
QY 1048 TGGGNTTAAACGCTGGATTGCTGCGGAGAGAGATCATGCGGAGTTCGACCTGG 1107
|||
Db 39972 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGCGCGGCTCGACAGTGGGAGCCGATGAGG 39913
|||
QY 1108 TATACTGGGGGCTTGACAAACAGCCTCTTTCGAGATAAACCGGATA---AAAAACGGAAAC 1164
|||
```

```
Db 39912 TCTATTGGGCAAGGAAGCCACCTTGGCTCGGCGATGAGCGTTACAGCGGTAAAGCGGATC 39853
|||
QY 1165 TTCAAGAACCTCTTGGCGCCACGACATGGGACTTATTTATGTCAATCTCTGAAGGCGCCG 1224
|||
Db 39852 TGGAGAAACCCGCTGGCGCGGTGAGATGGGCTGTACTAGTGAACCCGAGGCGCGCA 39793
|||
QY 1225 GTGGAACACAGATCTCTCTGGCTTCCGCGAAAGATATCAGGGAAGCTTTTTCACGTATGG 1284
|||
Db 39792 ACGCAACCCGACCCCATGCGCGCGGTGACATTCGCGAGACGCTTTCGCGCGCATGG 39733
|||
QY 1285 CCATGGATGATGAGGAGACTGTGGCCCTGTATCGCGGAGGCGCATACATTTTGGTAAAGCAC 1344
|||
Db 39732 CCATGAACGACGTGCAAAACAGCGCGCTGTATCGTGGCGGTTCACACTTTCGGTAAGACCC 39673
|||
QY 1345 ATGATGACGCTCTCTGAAATAATGTATTCGCGCAGGCGCTGTATGTTGTCACCTGTGAGG 1404
|||
Db 39672 ATGGCGCG---GCCCGCGCATCTGGTTCGCGCCCGGAACCGAGGCTGCTCCGCTGGAGC 39616
|||
QY 1405 AGCAGGACTGGGATGGAATAATAATGTGTACAGGAACGGCAAAATATACCATCACCA 1464
|||
Db 39615 AGATGGGCTTGGGCTGGAAGAGCTCGTATGGCACCGGAACCGGTAAGACGCGATCACCA 39556
|||
QY 1465 GTGGCCTTGAAGGAGCCTGTGTCGA 1488
|||
Db 39555 GCGGCATCGAGGTCTGTATGGACGA 39532
|||

RESULT 21
US-10-468-356-645/c
; Sequence 645, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 645
; LENGTH: 82993
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-645

Query Match 20.0%; Score 298.4; DB 18; Length 82993;
Best Local Similarity 62.4%; Pred. No. 8.3e-87;
Matches 502; Conservative 0; Mismatches 296; Indels 6; Gaps 2;

QY 688 ATCCCTGGGGGCTGATTTTGTATTCGATCCAGGATTTCAACAGCTGGATATGAGGCTC 747
|||
Db 40332 ACCCGATGGGTGGCGCGTTCGACTATGCCGGAGGTGCGGACCATCGACGTGACGCCC 40273
|||
QY 748 TGAATAAAGATATCAAGATTTTGTGCAAACTTCCAGGATTTGTCGCCCTCGGATTATG 807
|||
Db 40272 TGACGCGGACATCGAGGAGTGATGACCACTCGCAGCCGTGTGGCCCGGACTACG 40213
|||
QY 808 GTCAATTATGTCCTTTTATTCGATGCTGCGCAGGATTTGTCGCCCTCGGATTATG 867
|||
Db 40212 GCCACTACGGCGCTGTTTATCCGGATGGCGTGGCAGCTGCGGCACTTACCGCATCC 40153
|||
QY 868 ATGATGCGCGGAGGCGCCAGTGTGTGACAACTTTTGAACCGCTGAACAGCTGCC 927
|||
Db 40152 ACGACGCGCGCGCGCGCGGCGCATGACGCGGTTCGCGCGCTTAAACAGTGGC 40093
|||
QY 928 CGGATAACGTTAACTCGATATAAGCCCTCTTGCAGATAAACCGGATA---AAAAACGGAAAC 987
|||
Db 40092 CCGACNACCCAGCTTTGACNAGGCGCGCGGCTGTGTGCGGTCGAGNAGATACG 40033
|||
QY 988 GCTCCAGTATTTCTTGGGAGACCTGATGGTCCTGACTGGTAATGTTGCCCTTGAATCCA 1047
|||
Db 40032 GCAAGAAGCTCTCATGGCGGACCTGATTTGTTTTCGCGGCAACTGCGCGCTGGAATCGA 39973
|||
QY 1048 TGGGNTTAAACGCTGGATTGCTGCGGAGAGAGATCATGCGGAGTTCGACCTGG 1107
|||
Db 39972 TGGGCTTCAAGACGTTTCGGGTTTCGGCTTCGCGCGGCTCGACAGTGGGAGCCGATGAGG 40093
|||
QY 928 CGGATAACGTTAACTCGATATAAGCCCTCTTGCAGATAAACCGGATA---AAAAACGGAAAC 987
|||
```


Db 40092 CCGACAGCCAGCTTGGACAGGCGCCCGCTGCTGTGGCGGTCAAGAAAGTACG 40033
Qy 988 GCTCCAGTATTTCTGGGGAGACCTGATGGTCTGACTGGTAATGTTGCCCTTGAATCCA 1047
Db 40032 GCAAGAAGCTCTCATGGCGGACCTGATGTTTTCGCGGCAACTGCGCGCTGGAATCGA 39973
Qy 1048 TGGGATTTAAACGCTGGGATTTGCTGGCGGAAGAGATGACTGGAGTCCGACCTGG 1107
Db 39972 TGGGCTTCAAGACGTTGGGTTTCGGCTTCGGCCCGGTTCGACAGTGGGAGCCGATGAGG 39913
Qy 1108 TATAGTGGGGCTTGACAAACAAGCCTCTTGACAGATAACCGGATA- --AAAACGGGAAC 1164
Db 39912 TCTATGGGGCAAGAGACCACTGGCTCGGGATGACGTTACAGCGTTAAGCGGATC 39853
Qy 1165 TTCAGAAACCTTTCTGGCCGACGAGATGGACTTATTTATGTCAATCCTGAAAGCCCCG 1224
Db 39852 TGGAGAACCCCTGGCCCGGTGCAGATGGGCTGATCTACGTGAACCCGAGGGCCGA 39793
Qy 1225 GTGGAACACAGATCCTTGGCTTCGCGGAAGATATCAGGAAGCTTTTTCACGTATGG 1284
Db 39792 ACGGCAACCCGAGACCCATGCGCGCGGTTCGACATTCGCGAGACGTTTCGGCGCATGG 39733
Qy 1285 CCATGGATGATGAGAGACTGTGGCCCTGTATCGCGGAGGCGATACATTTTGTAAAGCAC 1344
Db 39732 CCATGAACGAGCTGGAACAGCGGCTGTATGCTCGCGGTTCACATTTTGGTAAGACC 39673
Qy 1345 ATGTGACAGCTCTCTGAAATAATGTATTTGGCGCAGGCGCTGTATGTCACCTGTGGAGG 1404
Db 39672 ATGGCGCG- --GCCCGCGATCTGGTTCGCGCCCGCAACCCGAGGCTGCTCGCTGGAGC 39616
Qy 1405 AGCAGGACTGGATGGAATAAATGATGTGTACAGGAACGGCAATATACCATCACCA 1464
Db 39615 AGATGGGCTTGGGCTTGAAGAGCTCGTATGGCACCGGAACCGGTAAGACGCGATCACCA 39556
Qy 1465 GTGGCTTGAAGGAGCTGTGTCGA 1488
Db 39555 CGCGATCGAGTCTGTATGGACGA 39532

RESULT 22
US-10-369-493-41662
; Sequence 41662, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41662
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: SPHINGOMONAS
US-10-369-493-41662

Query Match 20.0%; Score 297.6; DB 17; Length 2133;
Best Local Similarity 60.5%; Pred. No. 1.5e-87;
Matches 520; Conservative 0; Mismatches 319; Indels 21; Gaps 1;
Qy 648 CTGATTTAACTCTCTGAGATTACACAGCCCTGAAATCCCTGGGGGCTGATTTT 707
Db 25 CTCGATCTGACCGTCTCCACCAAGACCCGCGCGAGCCGATGGGCGGCGACTTC 84
Qy 708 GATTATGCCACCAGATTTCAACAGCTGGATATGGAGGCTCTGAAAAAAGATATCAAGAT 767

Db 85 GATTATGCGCGAATTTCCGGAAGCTCGATCTGGCGGGTCAAGCGCGACATCTTCGCG 144
Qy 768 TTGCTGACAACTTCCAGGATTTGCTGCCCTCGGATTTATGCTCATTTATGTCCTTTCTTT 827
Db 145 CTGATGACCGATTCGACGAGTATGCTGGCGCGGACTTCGSCCATTTAGCGCCCGCTGTT 204
Qy 828 ATTCTATGCTTGGCAGGTCGCGGAACATACAGACATATGATGCGCGGGAGGCGCC 887
Db 205 ATCCGATGGCGTGGCACAGCGCGGACCTTATCGCATTCGCGCAGCGGCGCGCGCGCG 264
Qy 888 AGTGTGCTCAGCAACGTTTTTGAACCGCTGAACAGCTGGCCGATAAAGTTAAATCTCGAT 947
Db 265 GGCTCGGATACGACGCTTCGCCCGCTCAACAGCTGGCCCGCAATTCGCAATCTCGAC 324
Qy 948 AAAGCCCGTGAATGCTGTGGCCAGTCAAGAAAAATAACGGCTCCAGTATTTCTCTGGGA 1007
Db 325 AAGCGCGCATGCTGCTGTGGCCGATCAAGCAGAAATATGGCCGGCGCTGTCTCTGGGCC 384
Qy 1008 GACCTGATGCTCTGACTGTAATGTTGCCCTTGAATCCATCGGATTTAAACCGTGGGA 1067
Db 385 GATCTGATGATCTCTCGCGGCAATTTGCGCGCTCGAATCGATGGGCTTCAAGACGCGCGC 444
Qy 1068 TTTCTGCGCGAAGAGATGACTGGAGTTCGA- --CTG 1106
Db 445 TTGCGCGCGCGCTGCGGATGTGTGGAGCGCGGCGCTATTTCGGCGAGCGGACCTG 504
Qy 1107 GTATATCTGGGGCGCTGACAACAGCCTCTTTGCAGATAACCGGGATAAAAAACGGAACTT 1166
Db 505 CGCGAATGGCTGCAGACCAAGCGCGCGCGCTATTTCGGCGAGCGGACCTG 564
Qy 1167 CAGAAACCTCTTTGCGCGCACGAGATGGACTTATTTATGTCATCTCTGAAAGGCCCGGT 1226
Db 565 TCGAACCCGCTGGCGGCGGTGCAGATGGGCTCTATCTACGTCAATCCGAAAGGCGCGAC 624
Qy 1227 GGAACACAGATCTCTGGCTTCGCGAAGATATCAGGAAGCTTTTTCAGCTATGGCC 1286
Db 625 GGCACCCCGATCCGCTGCTCGCGCGCACAGCATCCGCGAGACCTTCGCGCGCATGGCG 684
Qy 1287 ATGATGATGAGGAGACTGTGGCCCTGATCGCGGAGGAGATACATTTGTTAAAGCACAT 1346
Db 685 ATGAACGACGAGGAGACCGTCGCGCTGATCGCGCGGCGCATACCTTCGCGCAAGACCAT 744
Qy 1347 GGTGACCGCTCTCTGAAATAATGATTTGGCGCAGGCGCTGATGTGACCTGTGGAGAG 1406
Db 745 GGTGCGCGCGACCCAGCCAGCATATCGGGCGGAGCCGAAAGCGCGGCGCTGGAGATG 804
Qy 1407 CAGGACTGGGATGGAATAAATGTTGACAGGAACCGCAATATACCATCACCAT 1466
Db 805 CAGCGCTCGCTGGCGAACAGCTACGCAACCGCAATGCGCGGACATGCGCGGACATCACGAC 864
Qy 1467 GGCCTGGAAGGAGCTGTGTC 1486
Db 865 GGCCTGGAGGTCACTCGAC 884

RESULT 23
US-10-369-493-45100
; Sequence 45100, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21

```
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45100
; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Archaeoglobus fulgidus
US-10-369-493-45100

Query Match      20.0%; Score 297.2; DB 17; Length 2226;
Best Local Similarity 62.3%; Pred. No. 2e-87;
Matches 502; Conservative 0; Mismatches 298; Indels 6; Gaps 2;

Qy 687 AATCCCTGGGGGCTGATTTTGGATTATGCCACAGATTTCAACGACTGGATATGGAGCT 746
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
103 AATCCCTACGGGAGGATTACGATTACGTCAGGAGGTTGAGATCTCGATATCGATGCT 162
Qy 747 CTGAAAAAGATATCAAGATTGCTGCACACATCTCCAGAGATTGGTGGCCCTGGCGATTAT 806
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 GTCATCAGGGATTGAAGGAGCTGATGAGAGCTCTCAGGACTGGTGGCCGCTGACTTC 222
Qy 807 GGTCAATTATGCTCTCTTTTATTCGTATGCTTGGCACGCTGCGGAACATACAGGACA 866
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 GGCATTATACGCCCTTTGTTTATCCGCTTCGCTGGCATAGCGGGGAGCTACCGCAATT 282
Qy 867 TATGATGCCGGGAGGCGGCAGTGGTGTGAGCAACGTTTTTGAACCGCTGAAACAGCTGG 926
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 TTTGATGCGAGGTGGGCGAGGAGCGGAGCATCCGCTTTCCGCCCGCATAACTGG 342
Qy 927 CCGATTAACGTTAATCTGGATGAAGCCCGTGCATGCTGTGGCCAGTCAAGAAAAATAC 986
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 CCGACAACTAAACCTCGACAGCGGATAGGCTGCTCTGGCCGATAAAGAAAGTAC 402
Qy 987 GGCTCCAGTATTTCTGGGAGACTGATGCTGCTGACTGTGTAATGTTGCCCTTTGAATCC 1046
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
403 GGCAGAAAGCTGCTTGGGCGGATCTGATTAATCTTGGGCGACAGTTGCAATCGAGAC 462
Qy 1047 ATGGGATTTAAACCGCTGGGATTTCTGGCGGAAGAGAGATGACTGGAGTCCGACCTG 1106
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
463 ATGGGAGTGAAGCTTTTCGGCTTCGCACTGGCGAGGAGGACATCTTTGAGCCTGACGAG 522
Qy 1107 GTAT---ACTGGGGCTGACACACAGCTCTTGCAGATACCGGGATAAACCGGAAA 1163
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
523 AGCCCCGACTGGGGCGGAGGAGATGCTTACAGCAAGAGAGGGGAGAGAGAGAG 582
Qy 1164 CTTCAGAACTCTTTGGCGCCACGAGATGGACTTATTTATGTCAATCTCTGAAGGCC 1223
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
583 CTTGAGAGACCTTTTGGAGCTACCGAGATGGGCTGATTTACGTAAACCTTGAAGGCC 642
Qy 1224 GGTGAAACACAGATCTCTCGCTTCCGCGAAAGATATCAGGGAAGCTTTTTCAGTATG 1283
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
643 GCGGAAATCCAGATCCGCTTGGCTCGGCTCAGGAGATTAGGTTGCCCTTCGCGAGGATG 702
Qy 1284 GCCATGATGATGAGGAGACTGTGGCCTGATCGGGGAGGCGATACATTTGGTAAAGCA 1343
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
703 GGGATGAACGACGAGGAGACTGTAGCACTTATCGCGAGGCGATGCCCTTCGGAAGTGT 762
Qy 1344 CATGTGCGAGCTCTCTGAAATATGATTGCGCGAGGCGCTGATGGTGCACTGTGGAG 1403
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
763 CATGCGCTGACCGCTG---ATTATCTTGGCCAGACCCCAAGCTCTCTCCCATTTGAA 819
Qy 1404 GAGCAGGCACTGGGATGGAATAAATGTGGTACAGGAAACGCGCAATATACCATCAC 1463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
820 ATGACGGGCTCGGATGGAAGTACAACTACGCGAAGGAAAGGGCTCTGACACCTTCACT 879
Qy 1464 AGTGGCTGGAAGAGGCTTGTGCAC 1489
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
880 TCTGTCTTGAAGTTACTTGTGCGC 905
```

RESULT 24
US-10-282-122A-13926
; Sequence 13926, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13926
; LENGTH: 2244
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-13926
```

```
Query Match      19.7%; Score 293.6; DB 17; Length 2244;
Best Local Similarity 59.7%; Pred. No. 3.3e-86;
Matches 541; Conservative 0; Mismatches 344; Indels 21; Gaps 2;

Qy 602 AGCGCTGATAAAAAAGAGACTCAAAAATTTCTACTATCCAGAAACACTGGATTTAACTCC 661
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
93 AGCGCAGGCNACGGCAGCTCGAACCGGACTGGTGGCCCAATCAGCTGGACCTGAGCAT 152
Qy 662 TCTGAGATTACAGCCCTGAATCAAAATCCCTGGGGGGCTGATTTTGATTTATGCCACAG 721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
153 CCTGCATCGGCACCTCGCTCGCTGTCGATCCGATCCGCAAGGATTTCAACTACGGCGCAGGC 212
Qy 722 ATTTCACAGCTGGATATGGGCTCTGAAAAAAGATATCAAGATTTTCTGACAACTTC 781
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
213 GTTCGAGAGCTGACCTCGCGGGGTGAAGCGCGACTCCACGCGCTGATGACAGCTC 272
Qy 782 CCAGATTGTTGCCCTCGGATTTATGTTCAATTTATGTTCTCTTTTATTCGTATGGCTTG 841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
273 GCAGACTGTTGGCGCGCGATTTTCGGCCACTACGGCGGCCTGTTTCATCCGATGGCATG 332
Qy 842 GCACGGTCCGGAACATACAGGACATATGATGGCGGGGAGCGCAGTGTGTGTCAGCA 901
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
333 GCACAGCGCGGCACGTACCGCACCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 392
Qy 902 ACGTTTGTAAACCGCTGAACAGCTCGCGCGATAAGCTTAACTCTGGATAAAGCCCGCTCGATT 961
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
393 GCGCTTCGCGCGCTCAACAGCTGGCCGACACAGCGNACCTCCACAAAGGCGCGCGCGCT 452
```

QY 962 GCTGTGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGGTCT 1021
Db 453 GCTGTGCCAGTCAAGCAAGTACGGCCGCGCATCTGTGGCCGACCTGCTGATCCT 512
QY 1022 GACTGGTAATGTTGCCCTTGAATCCATGGGATTTAAACCGTGGGATTTGCTGGCGAAG 1081
Db 513 GACGGCAACGTCGCGCTCGAATCGATGGGCTTCAAGACCTTCGGCTTCGGGGCGCGC 572
QY 1082 AGAAGATGACTGGAGTGGACCTGGTATATCTGGGGCTGACAAAGCCCTCTTGAGA 1141
Db 573 CGCGGACACGTTGGAGCCCGAGGACGCTACTGGGGCTCGAAAAAGATCTGGCTGAACT 632
QY 1142 TAACCGGGATAA---ACGGGAACCTTCAGAAACCTCTTGGCGC 1183
Db 633 GAGCGGGCCCGCAACAGCCGCTATTGGGGACCGCCAGCTCGAGAACCCGCTCGCCGC 692
QY 1184 CACGAGATGGGACTTATTTATGTCAATCTCTGAAGCCCGCTGGAAAAACAGATCCTCT 1243
Db 693 CGTCAGATGGGCTCATCTACGTGAATCCGGAAGCCCGACGCAATCCGATCCGGT 752
QY 1244 GGCTTCGCGAAGATATCAGGAAGCTTTTTCAGTATGGCCATGGATGATGAGGAGAC 1303
Db 753 CGCGCGCGCGGACATTCGTGACACCTTCGCGCGCATGGCGATGAACGACGAGAGAC 812
QY 1304 TGTGCCCCGATCGCGGGAGGCGATACATTTGTGAAGCACATGTCAGCGTCTCTGA 1363
Db 813 GGTGCGGCTGATCGCGGGCGCCACAGCTTCGGAAGACGACCGCGGGCGGCGCGTC 872
QY 1364 AAATGTATTGGCGCAGGCGCTGATGTGTGACCTGTGGAGGAGCAGGAGCTGGGATGAA 1423
Db 873 GAA---CGTCGGCGCCGAGCGGAGGCGCGGGCATCGAAGCGAGGCGCTCGACTGNA 929
QY 1424 AAATAATGTGTACAGAAACGCGCAATATACCATCACAGTGGCTTGAAGAGCGCTG 1483
Db 930 GAGCGGTACCGACGCGGCGAAGGCGCGGACGCGATCAGCAGCGGCTCGAAGTCAAGTG 989
QY 1484 GTGAC 1489
Db 990 GACGAC 995

RESULT 25

US-10-282-122A-38893
; Sequence 38893, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; FILE OF INVENTION: Identification of Essential Genes in Microorganisms
; TITLE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38893
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Salmonella typhimurium
US-10-282-122A-38893

Query Match 19.5%; Score 289.8; DB 17; Length 2181;
Best Local Similarity 60.2%; Pred. No. 6e-85;
Matches 499; Conservative 0; Mismatches 327; Indels 3; Gaps 1;

QY 659 TCCTCTCAGATTACACAGCCCTGAATCAAAATCCCTGGGGGCTCATTTTGTATTATGCCAC 718
Db 132 TCCTTTGAATCAACATTCCCAACCGTCTTAACCCGCTGGGTGAAGACTTTGACTACCGCA 191
QY 719 CAGATTTCAACAGCTGGATATGAGGCTCTGAAAAAAGATATCAAGATTTGCTGACAA 778
Db 192 AGAGTTTAGCAGTTAGACTACTCCGCGCTGAAAGGGATCTCAAGCGCTGCTGACCGA 251
QY 779 TTCCAGGATTTGGCTCCCTGGGATTTATGTCATTATGCTCTTCTTTTATTCGTATGGC 838
Db 252 TTCAACACCGTGGTGGCCCTGACTGGGGCAGCTATGTCGGTTTGTATTTCGCATGGC 311
QY 839 TTGGCAGCGTCCCGAAACATACAGGACATATGATGGCCGGGAGCGCCAGTGTGTGCA 898
Db 312 CTGGCATGGCTGGCACCTACCGTTCTATTGATGGTCTGGCGCGCGGTCTGGTCA 371
QY 899 GCAACGTTTTGAACCGCTGAACAGCTGGCCGATACAGTTAATCTGATTAAGCCCGTCG 958
Db 372 ACAGCGTTTTGGCGCGCTTAACCTCTGGCGGATACAGTCAAGCTGATGAAGCGCGTCG 431
QY 959 ATTGCTGTGGCCAGTCAAGAAAAAATACGGCTCCAGTATTTCTGGGGAGACCTGATGGT 1018
Db 432 TTTGTTGTGGCCGATTAAAGCAGAAATATGGCCAGAAAAATTTCTGGGCGCACCTGTTAT 491
QY 1019 CCTGACTGTGTAATGTTGCCCTTGAATCCATGGGATTTAAAAACGCTGGGATTTGCTGGCG 1078
Db 492 TCTGGCGGTTAACTGGCGCTGGAAAACTCCGCGCTTCGCTACCTTCGGTTTCGGCGCGCG 551
QY 1079 AAGAGAGATGACTGGGAGTCCGACCTGGTATATCTGGGGGCTGACAAAGCTCTTGC 1138
Db 552 GCGTGAAGATGTCTGGGAACCGGATCTGGATGTGAACCTGGGGCGATGAAAAAGCTGGTT 611
QY 1139 AGATACCGGGATAAAAACCGGAAACTTTCAGAAACCTTTGCCGCCACGAGATGGGACT 1198
Db 612 GACTCACCGACACCTTGAAGCGCTGGCAAAAGCGCGCTGGGGCGACCCAGATGGGCT 671
QY 1199 TATTATGTCAATCTGAAGGCCCGCGTGGAAAAACAGATCTCTGGCTTCGCGGAAAGA 1258
Db 672 TATCTAGTTTAAACCGGAGGCGCGGATCACAGGGCGAACCACTTTCTCGCGCGCGCG 731
QY 1259 TATCAGGGAAGCTTTTTCAGTATGGCCATGGATGATGAGAGACTGTGGCCCTGATCGC 1318
Db 732 TATTGCGCTACCTTTGGCAATATGGGATGAACGACGAAGAGACCGCTGGCGTTGATCGC 791
QY 1319 GGGAGGCGCATACATTTGGTAAAGCATGTGACGCTCTCTGAAAAATGTTATTTGGCG 1378
Db 792 TGGCGGGCATACCTCTGGTAAAAACCCACGG---CGCGGACGCGCATCCCATGTAGGGC 848
QY 1379 AGGCGCTGATGGTGCACTCTGGAGGAGAGGAGGACTGGGATGAAAAATAAATGTTGGTAC 1438
Db 849 CGATCCGGAAGCGCGCGGATTAAGCGCAAGGCTTAGGTTGGCCAGCAGCTATGTTAG 908

QY 1439 AGGAACGGCAAAATATACCATCACCTAGTGGCTGGAGGAGCCTGGTCTG 1487
Db 909 TGGGTTGGCGCGGATGCTATCACCTCCGGGCTGGAAAGTGGTCTGGACG 957

Search completed: March 12, 2005, 00:25:13
Job time : 7445.72 secs